

Alkalmazott matematikai lapok

13

317.471

27/2010

2010/1

317.471

27/2010

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI TUDOMÁNYOK
OSZTÁLYÁNAK KÖZLEMÉNYEI

27.

KÖTET

ALKALMAZOTT MATEMATIKAI LAPOK

A MAGYAR TUDOMÁNYOS AKADÉMIA MATEMATIKAI TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

ALAPÍTOTTÁK

KALMÁR LÁSZLÓ, TANDORI KÁROLY, PRÉKOPA ANDRÁS, ARATÓ MÁTYÁS

FŐSZERKESZTŐ

PÁLES ZSOLT

FŐSZERKESZTŐ-HELYETTESEK

BENCZÚR ANDRÁS, SZÁNTAI TAMÁS

FELELŐS SZERKESZTŐ

VIZVÁRI BÉLA

TECHNIKAI SZERKESZTŐ

KOVÁCS GERGELY

A SZERKESZTŐBIZOTTSÁG TAGJAI

Arató Mátyás, Csirik János, Csiszár Imre, Demetrovics János, Ésik Zoltán, Frank András, Fritz József, Galántai Aurél, Garay Barna, Gécseg Ferenc, Gerencsér László, Györfi László, Györi István, Hatvani László, Heppes Aladár, Iványi Antal, Járai Antal, Kátai Imre, Katona Gyula, Komáromi Éva, Komlósi Sándor, Kovács Margit, Krisztin Tibor, Lovász László, Maros István, Michaletzky György, Pap Gyula, Prékopa András, Recski András, Rónyai Lajos, Schipp Ferenc, Stoyan Gisbert, Szeidl László, Tusnády Gábor, Varga László

KÜLSŐ TAGOK:

Csendes Tibor, Fazekas Gábor, Fazekas István, Forgó Ferenc, Friedler Ferenc, Fülöp Zoltán, Kormos János, Maksa Gyula, Racskó Péter, Tallos Péter, Temesi József

27. kötet 1. szám

Szerkesztőség és kiadóhivatal: 1027 Budapest, Fő u. 68.

Az Alkalmazott Matematikai Lapok változó terjedelmű füzetekben jelenik meg, és olyan eredeti tudományos cikkeket publikál, amelyek a gyakorlatban, vagy más tudományokban közvetlenül felhasználható új matematikai eredményt tartalmaznak, illetve már ismert, de színvonalas matematikai apparátus újszerű és jelentős alkalmazását mutatják be. A folyóirat közöl cikk formájában megírt, új tudományos eredménynek számító programokat, és olyan, külföldi folyóiratban már publikált dolgozatokat, amelyek magyar nyelven történő megjelentetése elősegítheti az elért eredmények minél előbbi, széles körű hazai felhasználását. A szerkesztőbizottság bizonyos időnként lehetővé kívánja tenni, hogy a legjobb cikkek nemzetközi folyóiratok különszámaként angol nyelven is megjelenhessenek.

A folyóirat feladata a Magyar Tudományos Akadémia III. (Matematikai) Osztályának munkájára vonatkozó közlemények, könyvismertetések stb. publikálása is.

A kéziratok a főszerkesztőhöz, vagy a szerkesztőbizottság bármely tagjához beküldhetők. A főszerkesztő címe:

Páles Zsolt, főszerkesztő

1027 Budapest, Fő u. 68.

A folyóirat e-mail címe: aml@math.elte.hu

Közlésre el nem fogadott kéziratokat a szerkesztőség lehetőleg visszajuttat a szerzőhöz, de a beküldött kéziratok megőrzéséért vagy továbbításáért felelősséget nem vállal.

Az Alkalmazott Matematikai Lapok előfizetési ára évfolyamonként 1200 forint. Megrendelések a szerkesztőség címen lehetségesek.

A Magyar Tudományos Akadémia III. (Matematikai) Osztálya a következő idegen nyelvű folyóiratokat adja ki:

1. Acta Mathematica Hungarica,
2. Studia Scientiarum Mathematicarum Hungarica.

SZINTÁTMETSZÉSI PROBLÉMA ÉS ÁLTALÁNOSÍTÁSA A SPARRE
ANDERSEN-MODELLBEN

MIHÁLYKÓNÉ ORBÁN ÉVA, MIHÁLYKÓ CSABA, LAKATOS G. BÉLA

A biztosítási matematika fő kérdése a tönkremenés valószínűségének, a tönkremenési idő eloszlásának meghatározása, és kevés figyelmet fordítanak a pénztárban levő pénzmennyiség legnagyobb értékének vizsgálatára. Mi most arra a kérdésre keressük a választ, hogy a biztosítási matematikában Sparre Andersen-modellként ismert modell esetén vajon meghalad-e valaha egy bizonyos szintet a pénztárban levő pénzmennyiség, és ha igen, akkor mikor. Általánosabban egy olyan függvény elemzését végezzük el dolgozatunkban, amely speciális esetként adja meg a pénzmennyiség-szint elérésének valószínűségét, valamint a szintátmetszési idő várható értékét.

1. Bevezetés

A biztosítási matematikában alkalmazott modellek esetén a központi kérdés, hogy véges vagy végtelen időintervallumon vizsgálva a folyamatot, vajon tönkremegy-e a biztosítótársaság, illetve ez az esemény mikor fog megtörténni. Ha más kérdések felvetődnek, akkor is általában a tönkremenés időpontjáig foglalkoznak a kérdéssel.

Mivel vannak olyan stratégiák, amelyben bizonyos pénzmennyiség elérésekor a befizetendő pénzmennyiséget csökkentik [1], vagy más esetekben osztalékot fizetnek egy bizonyos nyereség elérése után [4], ezért érdekes lehet annak vizsgálata, hogy vajon meghalad-e egy bizonyos szintet valaha is a pénztárban levő pénzmennyiség, és ha ez megtörténik, vajon mikorra várható. Vagyis milyen valószínűséggel, illetve mikor számíthatunk kedvezményekre. Mi most ezzel a kérdéssel olyan tekintetben foglalkozunk, hogy közben nem nézzük, vajon a pénztárban levő pénzmennyiség negatívvá vált-e eközben. Azért is érdekes lehet ez a kérdés, mert a tároló modellek vizsgálata során ennek a kérdésnek a megfelelője a szükséges kezdő anyagmennyiség meghatározása, amelynél nem releváns kérdés, hogy az anyagelfogyás előtt meghaladt-e egy bizonyos szintet a tárolóban levő anyag mennyisége [8, 9].

Ugyanez a kérdés vetődik fel abban az esetben, amikor a pénztárból a kifizetés járadék formájában történik, és a befizetések történnek véletlen időpontokban és véletlen nagyságban, és arra a kérdésre várunk választ, hogy vajon adott kezdő-tőke mellett milyen valószínűséggel megy tönkre a biztosítótársaság, illetve mikor

történik ez a tönkremenés. A kérdést megfogalmazza Grandell könyvében [6], és a tönkremenési probléma megoldását speciális esetben kapcsolatba hozza a klasszikus rizikófolyamat tönkremenési problémájának megoldásával.

A hagyományos rizikófolyamat és a pozitív ugrásokkal történő, általunk vizsgált folyamat tönkremenési egyenleteinek összekapcsolása általánosabb esetben történt Mazza és Rulliere cikkében [7]. A pozitív ugrások esetével foglalkozik, de csak a tönkremenési valószínűség megadását tárgyalja Dong és Wang Erlang(n) eloszlású káresemények közt eltelt idők esetén [2, 3]-ban.

A megfogalmazott célok érdekében végzett vizsgálatok során reflektorfénybe kerültek olyan függvények, amelyek elemzése segíti az elsődleges célok elérését, és választ adhatnak a gyakorlati problémákra. Egy ilyen függvénnyel foglalkozunk jelen dolgozatunkban.

2. A vizsgált modell

Tekintsük a biztosítási matematikában gyakran használt, Sparre Andersen-modellként ismert modellt, azaz legyen $t_0 = 0$, valamint legyenek t_i ($i = 1, 2, 3 \dots$) független, nemnegatív értékű azonos eloszlású valószínűségi változók. A biztosítási terminológiában t_i adja meg az $i - 1$ -edik és az i -edik kárkifizetések közt eltelt időt. Jelöljük közös eloszlásfüggvényüket $F(t)$ -vel, sűrűségfüggvényüket $f(t)$ -vel, μ_F -fel a közös (véges) várható értéküket és σ_F -fel a közös (véges) szórásukat. Jelölje $N(t)$ a t ideig történő káresemények számát. Az i -edik káresemény során az Y_i valószínűségi változó adja meg a kifizetendő pénzmennyiséget. Az Y_i valószínűségi változóról ugyancsak feltételezzük, hogy nemnegatív értékűek, egymástól függetlenek, és azonos eloszlásúak $G(y)$ eloszlásfüggvénnyel, $g(y)$ sűrűségfüggvénnyel, μ_G véges várható értékkel és σ_G véges szórással. Feltételezzük továbbá, hogy az $N(t)$ kárszám- folyamat és Y_i egymástól függetlenek. A befizetések folyamatosan érkeznek állandó c intenzitással. Jelöljük z_0 -al a kezdőtőkét. Ahhoz, hogy a pénztárban levő pénzmennyiség ne haladja meg a $z_1 \geq z_0$ szintet, az szükséges, hogy

$$z_1 \geq z_0 - \sum_{i=1}^{N(t)} Y_i + ct,$$

azaz a $z_1 - z_0 \geq ct - \sum_{i=1}^{N(t)} Y_i$ egyenlőtlenség teljesüljön minden nemnegatív t értékre. z -vel jelölve a $z_1 - z_0$ nemnegatív különbséget, vizsgáljuk a

$$z + \sum_{i=1}^{N(t)} Y_i - ct \geq 0$$

egyenlőtlenség teljesülését, ami azt jelenti, hogy a pénztárban levő pénzmennyiség növekménye nem haladja meg a z értéket.

Definiáljuk az $R : \mathbb{R}_0^+ \rightarrow \mathbb{R}$ függvényt az alábbi módon:

$$R(z) = P \left(\left\{ 0 \leq z + \sum_{i=1}^{N(t)} Y_i - ct, \quad \forall t : 0 \leq t \right\} \right),$$

továbbá vezessük be az alábbi függvényeket:

Legyen $V(t) = z - ct + \sum_{i=1}^{N(t)} Y_i$ és

$$T_V = \begin{cases} \inf \{t : V(t) < 0\} \\ \infty, & \text{ha } V(t) \geq 0 \quad \forall t \geq 0 \end{cases}$$

továbbá legyen $\phi : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \rightarrow \mathbb{R}$ az alábbi módon definiálva:

$$\phi(z, \delta) = E(e^{-\delta T_V} 1_{T_V < \infty} | V(0) = z), \quad z \geq 0, \quad \delta \geq 0.$$

Ez utóbbi függvény a szintátlépési idő sűrűségfüggvényének Laplace-transzformáltja. Cikkünkben a $\phi(z, \delta)$ függvénnyel kapcsolatos eredményeket kívánjuk ismertetni. Megemlítjük, hogy a korábban említett tönkremenési probléma esetén az imént definiált $\phi(z, \delta)$ -nak megfelelő függvényt Gerber és Shiu vezették be [5].

3. A $\phi(z, \delta)$ függvény általános tulajdonságai

A $\phi(z, \delta)$ függvénnyel kapcsolatban könnyen bizonyíthatjuk az alábbi állítást.

3.1. ÁLLÍTÁS. $\phi(z, \delta)$ korlátos, a második változójában monoton fogyó függvény.

Bizonyítás. Mivel $0 \leq e^{-\delta T_V} \leq 1$, és $0 \leq 1_A \leq 1$, ezért szorzatuk is 0 és 1 közé esik, tehát a szorzat várható értéke is 0 és 1 közé esik minden $z \geq 0$ -ra és $\delta \geq 0$ -ra. Ha δ nő, akkor $e^{-\delta T_V}$ szigorúan monoton fogy. \square

Megjegyezzük, hogy a $\phi(z, \delta)$ függvény a $\delta = 0$ paraméter esetén éppen $1 - R(z)$ -t adja meg, hiszen ha $\delta = 0$, akkor

$$\phi(z, 0) = E(1_{T_V < \infty}) = P(T_V < \infty) = 1 - R(z),$$

valamint a paraméterben folytonos, sőt akárhányszor deriválható ($\delta \geq 0$). Továbbá igaz, hogy

$$(-1)^k \frac{\partial^k \phi(z, \delta)}{\partial \delta^k} \Big|_{\delta=0} = E(T_V^k \cdot 1_{T_V < \infty}),$$

és

$$\lim_{\delta \rightarrow 0+} \phi(z, \delta) = \phi(z, 0) = 1 - R(z).$$

3.2. ÁLLÍTÁS. $\phi(z, \delta)$ kielégíti az alábbi integrálegyenletet, ha $z \geq 0$ és $\delta \geq 0$:

$$\phi(z, \delta) = \int_0^\infty \int_0^{\frac{z}{c}} e^{-\delta t} \phi(z + y - ct, \delta) f(t) g(y) dt dy + \int_{\frac{z}{c}}^\infty e^{-\delta \frac{z}{c}} f(t) dt. \quad (1)$$

Bizonyítás. z jelöli a beállított szintet, T_V pedig a szint elérési idejét, t_1 az első kifizetés idejét és Y_1 az első alkalommal kifizetett összeget. Ekkor

$$\begin{aligned} \phi(z, \delta) &= E(e^{-\delta T_V} 1_{T_V < \infty}) = E(E(e^{-\delta T_V} 1_{T_V < \infty} | t_1 = t, Y_1 = y)) = \\ &= \int_0^\infty \int_0^\infty E(e^{-\delta T_V} 1_{T_V < \infty} | t_1 = t, Y_1 = y) f(t) g(y) dt dy. \end{aligned} \quad (2)$$

Két eset lehetséges. Ha $t_1 \geq \frac{z}{c}$, akkor $E(e^{-\delta T_V} 1_{T_V < \infty} | t_1 = t) = e^{-\frac{z}{c}\delta}$, hiszen ekkor biztosan a $\frac{z}{c}$ időpontban éri el a szintet a pénznövekmény.

Ha $t_1 = t < \frac{z}{c}$, akkor a szint elérési idejét t_1 -hez viszonyíthatjuk, továbbá a t_1 időpontban a pénztárban levő pénz $z + y - ct$. Mivel

$$E(e^{-\delta T_V} 1_{T_V < \infty} | t_1 = t) = e^{-\delta t} \phi(z + y - ct, \delta),$$

ebből kifolyólag (2) a következővel lesz egyenlő:

$$\int_0^\infty \int_0^{\frac{z}{c}} e^{-\delta t} \phi(z + y - ct, \delta) f(t) g(y) dt dy + \int_0^\infty \int_{\frac{z}{c}}^\infty e^{-\delta \frac{z}{c}} f(t) g(y) dt dy.$$

A második integrált kiintegrálva y szerint éppen a kívánt állítást kapjuk. \square

3.3. ÁLLÍTÁS. Az előbbi egyenlet $\delta = 0$ esetén némi átalakítás után az alábbi egyenletet adja:

$$R(z) = \int_0^\infty \int_0^{\frac{z}{c}} R(z - c\tau + y) f(\tau) g(y) d\tau dy, \quad (3)$$

amely megegyezik a [8] publikációban közölt 3.2. állítás (7) egyenletével.

3.4. ÁLLÍTÁS. A (1) egyenletnek minden rögzített $\delta > 0$ esetén egyértelmű megoldása van az R_0^+ -on korlátos függvények körében.

Bizonyítás. Tegyük fel, hogy (1)-nek két különböző megoldása is van, és vegyük a különbségüket, jelöljük ezt $\phi_k(z, \delta)$ -val. Ekkor $\phi_k(z, \delta)$ kielégíti a

$$\phi_k(z, \delta) = \int_0^\infty \int_0^{\frac{z}{c}} e^{-\delta t} \phi_k(z + y - ct, \delta) f(t) g(y) dt dy$$

egyenletet.

Legyen $\|\phi_k(\cdot, \delta)\| = \sup_{z \in R_0^+} |\phi_k(z, \delta)|$. Mivel mindkét megoldás korlátos, ezért a különbségük maximum normája véges. Továbbá

$$|\phi_k(z, \delta)| \leq \int_0^\infty \int_0^{\frac{z}{c}} e^{-\delta t} |\phi_k(z + y - ct, \delta)| f(t) g(y) dt dy \leq \|\phi_k(\cdot, \delta)\| \int_0^\infty e^{-\delta t} f(t) dt$$

mindenz ≥ 0 esetén.

Ha $\delta > 0$ és ha $\|\phi_k(\cdot, \delta)\| \neq 0$ (azaz a különbség nem azonosan 0), akkor azt kapjuk, hogy $\|\phi_k(\cdot, \delta)\| \leq \|\phi_k(\cdot, \delta)\| \int_0^\infty e^{-\delta t} f(t) dt < \|\phi_k(\cdot, \delta)\|$, ami ellentmondás. Ez azt jelenti, hogy nem lehet két különböző megoldása (1)-nek, ha $\delta > 0$. \square

Megjegyezzük, hogy $\delta = 0$ esetén $\int_0^\infty e^{-\delta t} f(t) dt = \int_0^\infty f(t) dt = 1$, tehát az egyértelműség az előbbi gondolatmenet segítségével nem igazolható. Sőt, az egyértelműség nem is igaz, hiszen a (3) egyenletnek az azonosan 0 függvény is megoldása. A $\lim_{z \rightarrow \infty} R(z) = 1$ feltétel mellett azonban speciális esetekben sikerült bizonyítani a megoldás egyértelműségét, s ez a megoldás természetesen az $1 - \lim_{\delta \rightarrow 0^+} \phi(z, \delta)$ függvény [10].

4. A $\phi(z, \delta)$ függvény tulajdonságai Poisson kárszám-folyamat esetén

Vizsgáljuk meg azt az esetet, amikor a kárszám-folyamat Poisson-folyamat.

4.1. ÁLLÍTÁS. Ha a kárszám-folyamat Poisson-folyamat, azaz $f(t) = \lambda e^{-\lambda t}$ minden nemnegatív t -re, akkor $\phi(w, \delta)$ kielégíti az alábbi integrálegyenletet:

$$\phi(w, \delta) - \phi(0, \delta) = -\frac{\delta}{c} \int_0^w \phi(x, \delta) dx + \frac{\lambda}{c} \int_0^\infty (\phi(w + y, \delta) - \phi(y, \delta))(1 - G(y)) dy \quad (4)$$

és $\phi(0, \delta) = 1$ minden $\delta \geq 0$ esetén.

Bizonyítás. Induljunk ki (1)-ből, írjuk be $f(t)$ helyébe az exponenciális függvényt, és végezzük el az integrálást! Azt kapjuk, hogy

$$\phi(z, \delta) = \int_0^\infty \int_0^{\frac{z}{c}} \lambda e^{-(\delta+\lambda)t} \phi(z + y - ct, \delta) g(y) dt dy + e^{-(\lambda+\delta)\frac{z}{c}}.$$

Vezessük be a $\tau = z - ct$ új változót; az integrálos tagot tovább vizsgálva

$$\begin{aligned} & \int_0^\infty \int_z^0 \frac{-\lambda}{c} \lambda e^{-(\delta+\lambda)\frac{z}{c}} \phi(y + \tau, \delta) e^{(\delta+\lambda)\frac{z}{c}} g(y) d\tau dy = \\ & = \int_0^\infty \int_0^z \frac{\lambda}{c} e^{-(\delta+\lambda)\frac{z}{c}} \phi(y + \tau, \delta) e^{(\delta+\lambda)\frac{z}{c}} g(y) d\tau dy. \end{aligned}$$

Beszorozva $e^{(\delta+\lambda)\frac{z}{c}}$ -vel és deriválva z szerint azt kapjuk, hogy

$$e^{(\delta+\lambda)\frac{z}{c}} \left(\frac{\partial \phi(z, \delta)}{\partial z} + \frac{\delta + \lambda}{c} \phi(z, \delta) \right) = \int_0^\infty \frac{\lambda}{c} \phi(y + z, \delta) e^{(\delta+\lambda)\frac{z}{c}} g(y) dy. \quad (5)$$

Egyszerűsítve $e^{(\delta+\lambda)\frac{z}{c}}$ -vel

$$\frac{\partial \phi(z, \delta)}{\partial z} + \frac{\delta + \lambda}{c} \phi(z, \delta) = \int_0^\infty \frac{\lambda}{c} \phi(y + z, \delta) g(y) dy.$$

Átrendezve

$$\frac{\partial \phi(z, \delta)}{\partial z} = \int_0^\infty \frac{\lambda}{c} \phi(z + y, \delta) g(y) dy - \frac{\delta + \lambda}{c} \phi(z, \delta).$$

Ezt az egyenletet integrálva z szerint 0-tól w -ig

$$\begin{aligned} \phi(w, \delta) - \phi(0, \delta) &= \int_0^w \int_0^\infty \frac{\lambda}{c} \phi(y + z, \delta) g(y) dy dz - \int_0^w \frac{\delta + \lambda}{c} \phi(z, \delta) dz = \\ &= \int_0^\infty \int_0^w \frac{\lambda}{c} \phi(y + z, \delta) g(y) dz dy - \int_0^w \frac{\delta + \lambda}{c} \phi(z, \delta) dz. \end{aligned}$$

Bevezetve az $x = z + y$ új változót és a $B(w, \delta) = \int_0^w \phi(x, \delta) dx$ új függvényt, az előbbi átalakítás a következőképpen folytatható:

$$\begin{aligned} &\int_0^\infty \int_y^{w+y} \frac{\lambda}{c} \phi(x, \delta) g(y) dx dy - \int_0^w \frac{\delta + \lambda}{c} \phi(z, \delta) dz = \\ &= \int_0^\infty \frac{\lambda}{c} (B(w + y, \delta) - B(y, \delta)) g(y) dy - \frac{\delta + \lambda}{c} B(w, \delta) = \\ &= \left[-\frac{\lambda}{c} (B(w + y, \delta) - B(y, \delta)) (1 - G(y)) \right]_0^\infty + \\ &+ \int_0^\infty \frac{\lambda}{c} (\phi(w + y, \delta) - \phi(y, \delta)) (1 - G(y)) dy - \frac{\delta + \lambda}{c} B(w, \delta) = \\ &= \frac{\lambda}{c} B(w, \delta) + \int_0^\infty (\phi(w + y, \delta) - \phi(y, \delta)) (1 - G(y)) dy - \frac{\delta + \lambda}{c} B(w, \delta). \end{aligned}$$

Az utolsó egyenlőségénél felhasználtuk, hogy

$$B(w + y, \delta) - B(y, \delta) = \int_y^{w+y} \phi(z, \delta) dz \leq w$$

minden rögzített w esetén és $\lim_{y \rightarrow \infty} (1 - G(y)) = 0$. Összegezve, azt kaptuk, hogy

$$\phi(w, \delta) - \phi(0, \delta) = \frac{-\delta}{c} \int_0^w \phi(x, \delta) dx + \frac{\lambda}{c} \int_0^\infty (\phi(w + y, \delta) - \phi(y, \delta)) (1 - G(y)) dy.$$

$\phi(0, \delta) = 1$ könnyen látható például $\phi(z, \delta)$ definíciójából. □

Megjegyezzük, hogy az előbb végrehajtott átalakítások fordított sorrendben is elvégezhetők. Azaz kiindulva a (4) egyenletből megkaphatjuk az (1) egyenletet, ha a bizonyításban szereplő, (5)-t eredményező lépésben említett deriválás ellentétéként történő integrálás során figyelembe vesszük a $\phi(0, \delta) = 1$ kezdőfeltételt. Így a (4) egyenletnek is egyértelmű megoldása van a korlátos és folytonos függvények körében.

4.2. ÁLLÍTÁS. *Ha a kárszám-folyamat Poisson-folyamat, akkor a (4) egyenletnek van exponenciális, azaz $\phi(z, \delta) = e^{-k(\delta) \cdot z}$ alakú megoldása, ahol $k(\delta) > 0$, értéke függ δ értékétől, de rögzített δ esetén egyértelmű. Ez a $k(\delta) > 0$ érték $\delta > 0$ esetén egyértelmű pozitív megoldása a*

$$k(\delta) \cdot \left(c - \lambda \int_0^\infty e^{-k(\delta)y} (1 - G(y)) dy \right) = \delta \quad (6)$$

úgynevezett általánosított Lundberg-egyenletnek. Megjegyezzük, hogy mivel a korábbiak során bebizonyítottuk, hogy az egyenlet megoldása egyértelmű a korlátos függvények körében, ezért ez a megoldás.

Bizonyítás. Induljunk ki az előzőleg bizonyított

$$\phi(w, \delta) - 1 = -\frac{\delta}{c} \int_0^w \phi(x, \delta) dx + \frac{\lambda}{c} \int_0^\infty (\phi(w+y, \delta) - \phi(y, \delta)) (1 - G(y)) dy$$

összefüggésből, és keressük a megoldást $\phi(w, \delta) = e^{-k(\delta)w}$ alakban!

Vizsgáljuk meg, milyen összefüggést kell kielégítenie a kitevőnek! Amennyiben $k(\delta) = 0$, akkor az egyenlet $0 = -\frac{\delta}{c} \int_0^w 1 dx$, azaz $\delta = 0$ vagy $w = 0$, de mivel az egyenletnek minden $w \geq 0$ esetén teljesülni kell, ezért $\delta = 0$. Legyen $k(\delta) \neq 0$! $k(\delta) < 0$ esetén rögzített $\delta > 0$ -ra a $\phi(z, \delta)$ nem korlátos. Ha $k(\delta) > 0$, akkor

$$\begin{aligned} e^{-k(\delta)w} - 1 &= \\ &= -\frac{\delta}{c} \int_0^w e^{-k(\delta)x} dx + \frac{\lambda}{c} \int_0^\infty \left(e^{-k(\delta)(w+y)} - e^{-k(\delta)y} \right) (1 - G(y)) dy = \\ &= -\frac{\delta}{c} \frac{e^{-k(\delta)w} - 1}{k(\delta)} + \frac{\lambda}{c} \left(e^{-k(\delta)w} - 1 \right) \int_0^\infty e^{-k(\delta)y} (1 - G(y)) dy. \end{aligned} \quad (7)$$

Mivel a kifizetendő kármennyiség várható értéke véges, így $\int_0^\infty (1 - G(y)) dy$ véges, tehát $\int_0^\infty e^{-k(\delta)y} (1 - G(y)) dy$ is véges. Legyen $D_2 = \int_0^\infty e^{-k(\delta)y} (1 - G(y)) dy$.

Így (7) alapján

$$\left(e^{-k(\delta)w} - 1 \right) \left(1 - \frac{\delta}{ck(\delta)} - \frac{\lambda}{c} D_2 \right) = 0$$

teljesül, s lévén hogy az első tényező nem 0, tehát

$$1 - \frac{\delta}{ck(\delta)} - \frac{\lambda}{c} D_2 = 0,$$

amiből

$$ck(\delta) - \delta - \lambda k(\delta)D_2 = 0,$$

vagyis

$$k(\delta)(c - \lambda D_2) = \delta \quad (8)$$

Korábban megállapítottuk, hogy $k(\delta) = 0$ esetén $\delta = 0$, ami az $R(z)$ esete. Ha $\delta = 0$ és $k(\delta) \neq 0$, akkor $c - \lambda D_2 = 0$. Mivel $D_2 = \int_0^\infty e^{-k(\delta)y}(1 - G(y))dy$, ezért ez azt jelenti, hogy $1 = \frac{\lambda}{c} \int_0^\infty e^{-k(\delta)y}(1 - G(y))dy$.

Ezzel megmutattuk, hogy állításunk általánosítása annak a [8]-ban bizonyított állításnak, amely kimondta, hogy Poisson kárszám-folyamat esetén $R(z)$ exponenciális alakú megoldása $R(z) = 1 - e^{-\nu z}$, ahol ν a $\frac{\lambda}{c} \int_0^\infty e^{-\nu y}(1 - G(y))dy = 1$ egyenlet megoldása.

Ha $\delta \neq 0$, akkor $k(\delta) \neq 0$, továbbá legyen

$$H(k) = k \left(c - \lambda \int_0^\infty e^{-ky}(1 - G(y))dy \right).$$

A (8) egyenlet a $H(k(\delta)) = \delta$ alakban írható. Ha k nő, akkor D_2 csökken, ellentettje nő, így $c - \lambda D_2$ nő. Előjele attól függően negatív, vagy pozitív kicsi k értékre, hogy $c - \lambda \mu_G \geq 0$, vagy $c - \lambda \mu_G < 0$. Ha $c - \lambda \mu_G \geq 0$, akkor

$$c - \lambda \int_0^\infty e^{-ky}(1 - G(y))dy > 0,$$

minden $k > 0$ esetén, tehát $H(k)$ szigorúan monoton nő. De $H(0) = 0$, és $\lim_{k \rightarrow \infty} H(k) = \infty$, ezért létezik pontosan 1 olyan $k > 0$ érték, amire $H(k) = \delta$. Ha $c - \lambda \mu_G < 0$, akkor kicsi k értékekre

$$c - \lambda \int_0^\infty e^{-ky}(1 - G(y))dy < 0,$$

így annak k -szorosa is negatív. Így $H(k)$ először fogy, majd nőni kezd, sőt bizonyos k értéktől fogva

$$k(c - \lambda \int_0^\infty e^{-ky}(1 - G(y))dy) > 0,$$

s ilyen k értékekre $H(k)$ szigorúan monoton nő, valamint a határértéke ∞ , tehát most is létezik pontosan egy olyan $k > 0$ érték, amelyre $H(k) = \delta$. A lépések fordított sorrendben való elvégzésével ellenőrizhető, hogy ezzel a $k(\delta)$ kitevővel a $\phi(w, \delta) = e^{-k(\delta)w}$ kielégíti a (4) egyenletet.

Ezzel beláttuk, hogy a (4) integrálegyenletnek van exponenciális megoldása, és az exponenciális függvény kitevője a (6) egyenlet egyértelmű pozitív megoldása. \square

Megjegyezzük, hogy a 4.2. tétel eredménye megtalálható Mazza és Rulliere [7] publikációjában is, ők azonban más úton jutottak el hozzá. Nevezetesen a klasszikus tönkremenési probléma megoldását használták fel a szintátmetszési probléma megoldásához, és az általuk bebizonyított kapcsolat segítségével jutottak el ehhez a formulához.

Általában azonban nem igaz, hogy az (1) integrálegyenlet megoldása exponenciális alakú. Fennáll ugyanis a következő:

4.3. ÁLLÍTÁS. *Ha a kárigények nagysága exponenciális eloszlású valószínűségi változó, akkor az (1) integrálegyenlet megoldása csak Poisson kárszám-folyamat esetén exponenciális alakú.*

Bizonyítás. Tegyük fel, hogy a $\phi(z, \delta) = e^{-\nu(\delta)z}$ függvény ($\nu(\delta) > 0$ mellett) kielégíti a (1) egyenletet, azaz

$$e^{-\nu(\delta)z} = \int_0^\infty \int_0^{\frac{z}{c}} e^{-\delta t} e^{-\nu(\delta) \cdot (z+y-ct)} f(t) \frac{1}{\mu_G} e^{-\frac{1}{\mu_G} y} dt dy + \int_{\frac{z}{c}}^\infty e^{-\delta \frac{z}{c}} f(t) dt.$$

Kiintegrálva y szerint, beszorozva $e^{\nu(\delta)z}$ -vel, majd deriválva z szerint, kapjuk, hogy

$$0 = \frac{1}{c} \frac{1}{\mu_G} \frac{1}{\nu(\delta) + \frac{1}{\mu_G}} e^{-\delta \frac{z}{c}} e^{\nu(\delta)z} f\left(\frac{z}{c}\right) + \left(\nu(\delta) - \frac{\delta}{c}\right) e^{(\nu(\delta) - \frac{\delta}{c})z} \left(1 - F\left(\frac{z}{c}\right)\right) - e^{(\nu(\delta) - \frac{\delta}{c})z} \frac{1}{c} f\left(\frac{z}{c}\right).$$

Egyszerűsítve $e^{(\nu(\delta) - \frac{\delta}{c})z}$ -val, majd bevezetve az $\tilde{F}(x) = 1 - F\left(\frac{z}{c}\right)$ jelölést, láthatjuk, hogy

$$0 = -\frac{1}{\mu_G} \frac{1}{\nu(\delta) + \frac{1}{\mu_G}} \tilde{F}'(x) + c \left(\nu(\delta) - \frac{\delta}{c}\right) \tilde{F}(x) + \tilde{F}'(x),$$

amiből

$$\tilde{F}'(x) = -\frac{c \left(\nu(\delta) - \frac{\delta}{c}\right)}{\mu_G \nu(\delta)} \tilde{F}(x).$$

Ez egy lineáris differenciálegyenlet $\tilde{F}(x)$ -re, amelynek az $\tilde{F}(0) = 1$ kezdőfeltétel mellett egyértelmű exponenciális megoldása van, ez pedig a következő:

$$\tilde{F}(x) = e^{-\frac{c \left(\nu(\delta) - \frac{\delta}{c}\right)}{\mu_G \nu(\delta) + 1} x}.$$

Visszahelyettesítéssel megkapjuk, hogy

$$\lambda = \frac{c \left(\nu(\delta) - \frac{\delta}{c}\right)}{\mu_G \nu(\delta) + 1} = \frac{c\nu(\delta) - \delta}{\mu_G \nu(\delta) + 1}$$

paraméterű exponenciális eloszlású valószínűségi változók a kárigények közt eltelt időközök.

Ebből $\nu(\delta)$ kifejezhető, és a pozitív gyök

$$\nu_1(\delta) = \frac{\delta\mu_G + \lambda\mu_G - c + \sqrt{(\delta\mu_G + \lambda\mu_G - c)^2 + 4\delta c\mu_G}}{2c\mu_G} > 0.$$

□

Megjegyezzük, hogy $\delta = 0$ esetén $\lambda \frac{\mu_G \nu(\delta)}{\mu_G \nu(\delta) + 1} = c\nu(\delta)$ miatt $\lambda = c(\nu(0) + \frac{1}{\mu_G})$ paraméterű exponenciális eloszlású valószínűségi változók az egymást követő káresemények.

Ezzel egyúttal megadtuk egy speciális esetben a (4) egyenlet megoldását, nevezetesen:

4.4. ÁLLÍTÁS. Ha $N(t)$ Poisson-folyamat, valamint $G(y) = 1 - e^{-\frac{1}{\mu_G}y}$ ($y > 0$), azaz a kifizetett károk nagysága is exponenciális eloszlású valószínűségi változó, akkor a kitevő analitikusan megadható, nevezetesen

$$k(\delta) = \frac{-\left(\frac{c}{\mu_G} - \lambda - \delta\right) + \sqrt{\left(\frac{c}{\mu_G} - \lambda - \delta\right)^2 + \frac{4c\delta}{\mu_G}}}{2c} \quad (9)$$

és így

$$\phi(z, \delta) = e^{-\frac{-\left(\frac{c}{\mu_G} - \lambda - \delta\right) + \sqrt{\left(\frac{c}{\mu_G} - \lambda - \delta\right)^2 + \frac{4c\delta}{\mu_G}}}{2c} z}.$$

Megjegyezzük, hogy ez az eredmény megtalálható Mazza és Rulliere [7] publikációjában is.

A (4) egyenlet megoldásának alakját figyelembe véve és felhasználva azt, hogy

$$(-1)^k \frac{\partial^k \phi(z, \delta)}{\partial \delta^k} \Big|_{\delta=0} = E(T_V^k 1_{T_V < \infty}),$$

$k(\delta)$ ismeretében könnyen meghatározható $E(T_V^k 1_{T_V < \infty})$.

4.5. ÁLLÍTÁS. Speciálisan, z kezdőtőkét feltételezve, ha $N(t)$ Poisson-folyamat, akkor

$$E(T_V 1_{T_V < \infty}) = -\frac{\partial \phi(z, \delta)}{\partial \delta} \Big|_{\delta=0} = zk'(0)e^{-k(0)z},$$

ahol

$$k'(0) = \frac{1}{\lambda \int_0^\infty ye^{-k(0)y}(1 - G(y))dy}.$$

Bizonyítás. Induljunk ki a (6) összefüggésből, azaz a

$$k(\delta) \cdot \left(c - \lambda \int_0^\infty e^{-k(\delta)y}(1 - G(y))dy \right) - \delta = 0 \quad (10)$$

alakból. Az implicit függvény tétel alapján láthatjuk, hogy $k(\delta)$ differenciálható. (10)-et deriválva δ szerint

$$k'(\delta) \cdot \left(c - \lambda \int_0^\infty e^{-k(\delta)y} (1 - G(y)) dy \right) + \\ + k(\delta) \cdot k'(\delta) \cdot \lambda \int_0^\infty y e^{-k(\delta)y} (1 - G(y)) dy - 1 = 0.$$

Figyelembe véve, hogy $\delta = 0$, valamint hogy $c - \lambda \int_0^\infty e^{-k(0)y} (1 - G(y)) dy = 0$, kapjuk, hogy

$$k'(0) = \frac{1}{\lambda \int_0^\infty y e^{-k(0)y} (1 - G(y)) dy}.$$

□

5. A kitevő numerikus meghatározása

Tekintettel arra, hogy az (1), de még a (4) egyenletnek is csak speciális esetben ismert az analitikus megoldása, ezért az egyenlet megoldása érdekében szükség van numerikus megoldási módszer megkonstruálására. Cikkünkben csak a (4) egyenlet numerikus megoldási módszerével foglalkozunk.

A (4) egyenlet megoldásának numerikus közelítése szempontjából nagyon hasznos, hogy tudjuk, hogy a (4) egyenletnek exponenciális függvény a megoldása, s a megoldás kitevőjében szereplő együttható meghatározható a (6) egyenlet megoldásával.

A gyökkereső eljárás során a Newton-módszert alkalmazzuk, kiszámítva a $H(k) - \delta$ függvény és k szerinti deriváltja értékeit, vagy ha azokat nem tudjuk kiszámítani, azoknak közelítő értékeit. Ez utóbbi esetben a közelítő értékek kiszámításához szükségünk van improprius integrál közelítő meghatározására. Ehhez a következőképpen jutunk:

Az $x = ky$ helyettesítés után $\frac{1}{k} (1 - G(\frac{x}{k})) = \tilde{g}(x)$ jelöléssel kapjuk, hogy

$$\int_0^\infty e^{-ky} (1 - G(y)) dy = \int_0^\infty e^{-x} \tilde{g}(x) dx,$$

s a numerikus integrálás elvégzésére a $\tilde{g}(x)$ függvénnyel a Gauss-Laguerre kvadratura formula használható [9].

A $H(k) - \delta$ függvény deriváltja a Newton-módszer alkalmazásához szintén kell, így meg kell (analitikusan vagy közelítőleg) azt is határozni. A derivált a következő alakú:

$$H'(k) = - \int_0^\infty y e^{-ky} (1 - G(y)) dy. \quad (11)$$

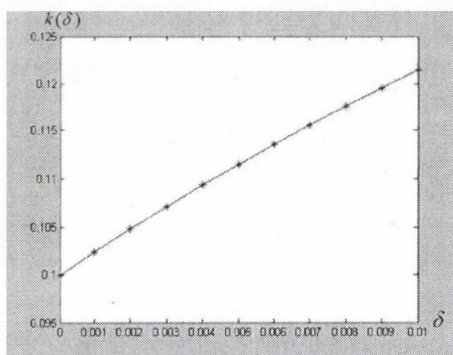
Az előbbi improprius integrál konvergencia, mivel

$$\begin{aligned} 0 &\leq \int_0^\infty y e^{-ky} (1 - G(y)) dy \leq \int_0^\infty y (1 - G(y)) dy = \\ &= \left(\left[\frac{y^2 (1 - G(y))}{2} \right]_0^\infty + \int_0^\infty \frac{y^2}{2} G'(y) dy \right) = \frac{1}{2} M_2, \end{aligned}$$

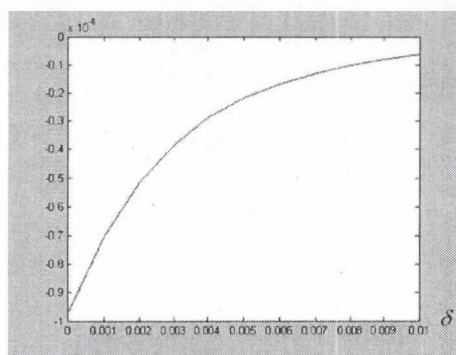
amennyiben M_2 jelöli Y_i $i = 1 \dots$ második momentumát.

A (11)-ben szereplő improprius integrál numerikus kiszámítására szintén használható a Gauss-Laguerre kvadratúra formula a $\tilde{h}(x) = \frac{x}{k^2} (1 - G(\frac{x}{k}))$ függvénnyel. Így a Newton-módszer alkalmazásával (6) gyöke numerikusan megadható, akár tudjuk a $H(k) - \delta$ függvény és k szerinti deriváltja értékeit pontosan számítani, akár nem. (10) és (11) alapján $k'(\delta)$ is megkapható numerikusan, így a szintátlépési idő várható értéke közelítőleg számolható.

A numerikus számolások során azt tapasztaltuk, hogy a tizenkilenced fokú Laguerre-polinomokat használva az improprius integrálok értékét nagy pontossággal megkaphatjuk. A tizenkilenced fokú polinom alkalmazásához szükséges kvadratúra súlyokat, valamint a polinom gyökhelyeit táblázatban találtuk meg [8].



(a)



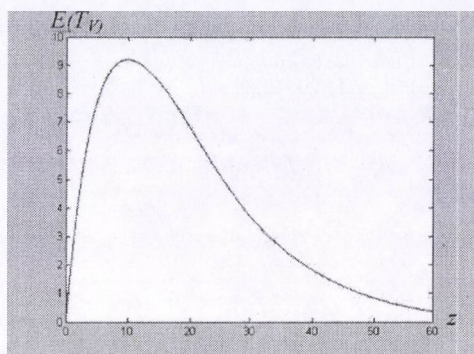
(b)

1. ábra. Az exponenciális megoldásfüggvény kitevőjének pontos (-) és numerikusan számolt (*) értékei (1(a)), és azok különbsége δ függvényében (1(b)).

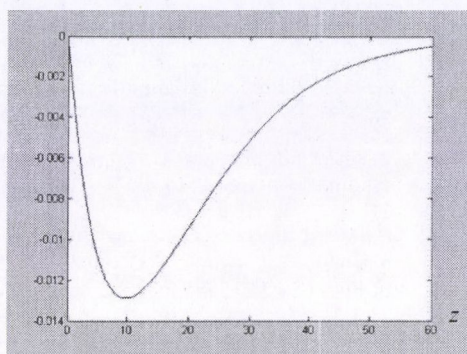
Végezetül a numerikus eljárás során kapott megoldás pontosságának illusztrálására mutatjuk az alábbi ábrákat. Legyen a kárszámot megadó folyamat Poisson-folyamat $\lambda = 1$ paraméterrel, azaz a folyamat során exponenciális eloszlású időközönként érkeznek a kárigények, egységnyi időközönként átlagosan egy, míg a kifizetett kár nagysága exponenciális eloszlású valószínűségi változó $\mu_G = 2,5$ (pénzegység) várható értékkel, legyen továbbá $c = 2$. Az 1(a) ábra az exponenciális megoldásfüggvény kitevőjének értékeit mutatja δ függvényében. A vonal a (9) alapján számolt pontos értéket, a * a numerikusan számolt értékeket mutatja. Az 1(b) ábra

pontos és a numerikusan számolt kitevők különbségét mutatja. Az eltérés nagyságrendje 10^{-6} .

A 2(a) ábra a szintátlépési idő várható értékét mutatja az átlépendő szint függvényében. A pontosan és a numerikusan számolt függvények az ábrán nem különböztethetők meg a nagyfokú egybeesés miatt. A numerikus és a pontos értékek különbségét a 2(b) ábrán vehetjük szemre. A különbség esetünkben századnyi nagyságrendű, és ott a legnagyobb, ahol maga a függvény is legnagyobb.



(a)



(b)

2. ábra. A szintátlépési idő várható értéke az átlépendő szint függvényében (2(a)), valamint a numerikusan számolt és a pontos várható érték különbsége (2(b)).

Első ránézésre kissé meglepő, hogy amint az ábrán látható, a szintátlépési idő várható értéke először nő, majd fogyni kezd. Ennek az az oka, hogy a szintátlépés valószínűsége a szint függvényében exponenciálisan csökken, tehát bizonyos szint esetén már nagyon kicsivé válik, így a véges várható érték a kicsi valószínűség miatt nem lesz nagy.

KÖSZÖNETNYILVÁNÍTÁS: A szerzők köszönetüket fejezik ki az ismeretlen bírálónak értékes, javító megjegyzéseikért.

Hivatkozások

- [1] BRATIYCHUK, M. S., DERFLA, D.: *On a modification of a the classical risk process*. Insurance: Mathematics and Economics **41** (2007), 156–162.
- [2] DONG, Y.: *Ruin probability for correlated negative risk sims model with Erlang processes*. Appl. Math. J. Chinese Univ. **24** (2009) 14–20.

- [3] DONG, Y., WANG, G.: *On a compounding assets model with positive jumps*. Applied Stochastic Models in Buisness and Industry **24**, (2008) 21–30.
- [4] GAO, Q. , WU, Y., ZHU, C. AND WEI, C.: *Ruin problems in risk models with dependent rates of interest*, Statistics and Probability Letters **77** (2007) 761–768.
- [5] GERBER, H. U., SHIU, E. S. W.: *On the time value of ruin*, North American Actuarial Journal, (1998) 248–78 .
- [6] GRANDSELL, J.: *Aspects of Risk Theory*. Springer-Verlag, New York, Berlin, London, Paris, Budapest, 1991.
- [7] MAZZA, C., RULLIERE D.: *A link between wave governed random motions and ruin processes*. Insurance: Mathematics and Economics **35** (2004) 205–222.
- [8] MIHÁLYKÓNÉ ORBÁN É., LAKATOS G. B., MIHÁLYKÓ CS.: *Tartálméretezési problémák vizsgálatának matematikai háttére sztochasztikus működési feltételek esetén*. Alkalmazott Matematikai Lapok **24** (2007) 277–301.
- [9] ORBÁN-MIHÁLYKÓ, É., LAKATOS, B. G.: *Intermediate storage in batch/semicontinuous processing systems under stochastic operational conditions*, Computers and Chemical Engineering, **28** (2004) 2493–2508.
- [10] ORBÁN-MIHÁLYKÓ, É., LAKATOS, B. G.: *On the advanced integral and differential equations of sizing procedure of storage devices*, Functional Differential Equations, **11** (2004) 121–131.
- [11] STROUD, A. H., SECREST, D.: *Gaussian quadrature formulas*, Pentice-Hall, London, 1966.

(Beérkezett: 2008. február 28.)

MIHÁLYKÓNÉ ORBÁN ÉVA

Pannon Egyetem

Matematika Tanszék

8200 Veszprém, Egyetem u. 10–12.

orbane@almos.uni-pannon.hu

MIHÁLYKÓ CSABA

Pannon Egyetem

Matematika Tanszék

8200 Veszprém, Egyetem u. 10–12.

mihalyko@almos.uni-pannon.hu

LAKATOS G. BÉLA

Pannon Egyetem

Folyamatmérnöki Tanszék

8200 Veszprém, Egyetem u. 10–12.

lakatos@fmt.uni-pannon.hu

LEVEL-CROSSING PROBLEM AND ITS GENERALIZATION
IN THE SPARRE ANDERSEN MODEL

ÉVA ORBÁN-MIHÁLYKÓNÉ, CSABA MIHÁLYKÓ, BÉLA G. LAKATOS

The main problem in insurance is the ruin probability and small attention turns to the maximum value of the surplus. In this paper we ask if, the surplus exceeds a given level or not, and when it does, supposing Sparre Andersen risk model. Moreover we define a generalization of the function giving the probability of the level crossing and we analyze it. During our investigation we do not bother if the surplus becomes negative.

BEJÁRÁSFÜGGETLEN SZEKVENCIÁLIS VÉKONYÍTÁS

KARDOS PÉTER, NÉMETH GÁBOR ÉS PALÁGYI KÁLMÁN

A váz mint régió-alapú alakjellemező szemléletesen definiált a préri-tűz hasonlattal: Az objektum határának minden pontját egyidejűleg meggyújtjuk és feltételezzük, hogy a tűz minden irányban egyenletes sebességgel terjed. Ekkor a vázat azon pontok alkotják, ahol az egymással versengő tűzfrontok találkoznak és kioltják egymást. A vékonyítás a tűzfront-terjedést modellezi diszkrét képtereken úgy, hogy a kapott vázközelítés topológiailag ekvivalens legyen a kiindulási objektummal. Az iteratív objektum-redukció egy lépésében csak az aktuális objektum határpontjai közül a törölhetőnek minősítettet távolítjuk el. Az eljárás terminál, ha az objektumon már nincs több törölhető pont.

A szekvenciális vékonyító eljárások kontúrkövetést alkalmaznak: bejárják az objektumok határpontjait és egyenként törlik a törlési feltételüknek eleget tevőket. Az így kapott vázak általában érzékenyek a határpontok bejárési sorrendjére. A jelen cikkben bemutatunk egy olyan 2-dimenziós szekvenciális vékonyító algoritmust, amely független a bejárési stratégiától, vagyis ugyanazt az eredményt adja a határpontok tetszőleges sorrendben történő vizsgálata mellett. A javasolt eljárásra bizonyítjuk a bejárásfüggetlenség és topológia-megőrzés tulajdonságokat.

Kulcsszavak: váz, vékonyítás, digitális topológia, topológia-megőrzés, bejárás-függetlenség.

1. Bevezetés

A vázat (*skeleton*) Blum vezette be mint a *középtengely-transzformáció* (*Medial Axis Transform, MAT*) eredményét [3]. A középtengely-transzformáció az objektum valamennyi pontjára megkeresi a hozzá legközelebbre eső határpontot(ka)t. Ha az eljárás valamely belső pontra egynél több legközelebbi határpontot talál, akkor azt a vázhoz tartozónak, vázpontnak minősíti. A vázat Blum egy szemléletes hasonlattal, a *préritűz terjedésével* illusztrálta: Ha a vizsgált objektum határának minden pontját egyidejűleg meggyújtjuk és feltételezzük, hogy a tűzfrontok minden irányba egyenletes sebességgel terjednek, akkor a váz azokból a pontokból áll, ahol az objektum belsejében az egymással versengő tűzfrontok találkoznak, kioltják egymást.

A váz mint régió-alapú alakleíró jellemző az elmúlt évtizedekben egyre fontosabb lett, kulcseleme számos, a képfeldolgozás és az alakfelismerés területén felmerült probléma megoldására javasolt módszernek [4].

A vázkijelölés (vagyis a matematikai váz egy alkalmas közelítése digitális bináris képeken) leggyakrabban alkalmazott módszere a front-terjedést modellező iteratív objektum redukció, a *vékonyítás* (*thinning*) [14]. A tűzfrontok terjedése – természeténél fogva – párhuzamos folyamat, így a legtöbb javasolt vékonyító algoritmus párhuzamos [5, 10, 14], vagyis az eljárás egy fázisában egyidejűleg távolítja el az aktuális objektum valamennyi törölhetőnek minősített határpontját. A szekvenciális vékonyító algoritmusok a kontúrkövetés technikáját alkalmazzák és egyenként távolítják el a törlési szabályukat kielégítő határpontokat [1, 9, 10].

A vázkijelölő eljárásokkal szemben támasztott két fő követelmény a topológia és az alak megőrzése. A digitális topológiában fontos fogalom az *egyszerű pont* (*simple point*): egy objektumpont akkor és csak akkor egyszerű, ha törlése topológia-megőrző redukció [8]. Az egyszerűség lokális tulajdonság, vagyis (az általánosan feltételezett topológiájú képeken) eldönthető a kérdéses pont 3×3 -as környezete alapján. Az alak-információ megőrzésére a vékonyító eljárások végpont-feltételeket használnak, vagyis a vonal-végpontok megőrzésével biztosítják azt, hogy egy tetszőleges, üreget nem tartalmazó objektum ne zsugorodjon össze egyetlen ponttá.

A párhuzamos vékonyító algoritmusok egyidejűleg több pontot törölnek, így nehéz a topológiai korrektség biztosítása és bizonyítása. Ráadásul a vékonyítás egy iterációs lépése nem oldható meg egyetlen, csupán a 3×3 -as lokális környezetet figyelő párhuzamos redukcióval [13]. A fentiek miatt a párhuzamos vékonyító algoritmusok vagy több fázisra, párhuzamos redukcióra bontanak fel egyetlen iterációs lépést (ahol az eredmény érzékeny a fázisok sorrendjére), vagy pedig a 3×3 -nál bővebb (általában nem szimmetrikus) környezettel adják meg a törölhető pontjaikat [5].

A szekvenciális vékonyító eljárások esetében a topológia-megőrzés könnyen garantálható, ha a törölhető pontok olyan egyszerű pontok, amelyek nem vonal-végpontok. Probléma viszont, hogy a javasolt eljárások vázközelítése függ a határpontok bejárasi sorrendjétől. A jelen cikkben egy olyan algoritmust javaslunk, mely ugyanazt az eredményt adja a határpontok tetszőleges bejárása mellett. Tudomásunk szerint ilyen eljárást még nem közöltek, a javasolt eljárás az első szekvenciális bejárásfüggetlen vékonyító algoritmus. Megjegyezzük, hogy Ranwez és Soille [12], valamint Iwanowski és Soille [7] eljárásai csupán olyan bejárásfüggetlen zsugorító algoritmusok, melyek vékonyításra csak a végpontok előzetes kijelölésével tehetők alkalmassá. Módszerünk hatékonyan implementálható és letisztult (számos algoritmusnál jóval kevesebb nemkívánatos ágot tartalmazó) vázat eredményez.

A cikk 2. fejezete ismerteti a digitális topológia legfontosabb fogalmait, a 3. fejezetben bemutatjuk a bejárásfüggetlenség garantálásában fontos kritikus pontpárokat és tulajdonságaikat. A javasolt algoritmus a 4. fejezetben található, melynek bejárásfüggetlenségét az 5. fejezetben bizonyítjuk. A 6. fejezetben néhány tesztképen bemutatjuk módszerünk eredményeit, összevetve azokat a közelmúltban közölt AK^2 algoritmus [2] által kivont vázakkal.

2. A digitális topológia alapfogalmai

Jelölje \mathbb{Z}^2 a 2-dimenziós sík egész koordinátájú pontjait, a továbbiakban pontokat. Legyen $x = (x_1, x_2)$ és $y = (y_1, y_2)$ két pont. A leggyakrabban használt *szomszédsági* (*adjacency*) relációkat az $\|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ (euklidészi) távolság segítségével adjuk meg. Az x és az y pontok *4-szomszédosak*, ha $\|x - y\| \leq 1$ és *8-szomszédosak*, ha $\|x - y\| \leq \sqrt{2}$. Jelölje $N_4(x)$ és $N_8(x)$ az x -szel 4-, illetve 8-szomszédos pontok halmazát, továbbá jelölje $N_4^*(x) = N_4(x) \setminus \{x\}$ és $N_8^*(x) = N_8(x) \setminus \{x\}$ a valódi 4-, illetve 8-szomszédokat.

A p pont *j-szomszédos* ($j = 4, 8$) a nem-üres $S \subseteq \mathbb{Z}^2$ ponthalmazzal, ha van olyan $s \in S$, hogy $s \in N_j(p)$. A különböző pontokból álló $\langle s_0, s_1, \dots, s_n \rangle$ sorozat $n \geq 0$ hosszú *j-út* ($j = 4, 8$) az s_0 pontból s_n -be az S ponthalmazban, ha a sorozat minden pontja S -beli és minden i -re ($1 \leq i \leq n$) s_i és s_{i-1} *j-szomszédosak*. (Meggjegyezzük, hogy $\langle s_0 \rangle$ egy 0-hosszú *j-út*.) Az $s_1 \in S$ és az $s_2 \in S$ pontok *j-összefüggők* ($j = 4, 8$) az S ponthalmazban, ha létezik *j-út* s_1 és s_2 között S -ben. Az S ponthalmaz *j-összefüggő* ($j = 4, 8$) az S' ponthalmazban ($S' \supseteq S$), ha S bármely két pontja *j-összefüggő* S' -ben. Könnyű belátni, hogy a *j-összefüggőségi* reláció – vagyis a reflexív és szimmetrikus *j-szomszédsági* reláció tranzitív lezártja – ekvivalencia-reláció valamennyi *j*-re ($j = 4, 8$). A *j-összefüggőségi* reláció tehát egy osztályozását adja meg egy tetszőleges ponthalmaznak, ahol az ekvivalencia-osztályokat *j-összefüggő komponenseknek* vagy *j-komponenseknek* (*j-component*) nevezzük.

Egy 2-dimenziós $(8, 4)$ *bináris digitális kép* (a továbbiakban $(8, 4)$ kép, vagy egyszerűen kép) a $(\mathbb{Z}^2, 8, 4, \mathbf{B})$ rendezett négyessel írható le [8], ahol a \mathbb{Z}^2 a képpontok halmaza; a $\mathbf{B} \subseteq \mathbb{Z}^2$ a fekete pontok halmaza, melynek pontjaihoz „1” értéket rendelünk; komplementere, a $\mathbb{Z}^2 \setminus \mathbf{B}$ a „0” értékű fehér pontok halmaza; a fekete pontokra a 8-összefüggőség érvényes, míg a fehérekre a 4-összefüggőséget tételezzük fel.

A 8-összefüggőségi ekvivalencia-relációval a fekete pontokat partícionáljuk. Az egy ekvivalenciaosztályba eső fekete pontok halmazát az adott kép *fekete komponensének* vagy *objektumának* nevezzük. Hasonlóképpen: a 4-összefüggőségi reláció a fehér pontok halmazát *fehér komponensekre* bontja fel. Egy $(8, 4)$ kép véges, ha a fekete pontok halmaza véges. Véges képen egyetlen végtelen fehér komponens található, amit *háttérnek* (*background*) nevezünk. A véges fehér komponens neve *üreg* (*cavity*). Egy véges kép reprezentálható egy P (véges) bináris tömbbel (bitmátrixszal), ahol valamennyi tömbön kívüli képpont értéke „0”.

A $(\mathbb{Z}^2, 8, 4, \mathbf{B})$ képen a $p \in \mathbf{B}$ fekete pont *határpont* (*border-point*), ha 4-szomszédos legalább egy fehér ponttal, azaz: $N_4(p) \setminus \mathbf{B} \neq \emptyset$. Jelölje $B(p)$ a p pont fekete 8-szomszédainak számát, vagyis az $N_8^*(p) \cap \mathbf{B}$ halmaz számosságát. A $p \in \mathbf{B}$ fekete pont *izolált* (*isolated*), ha $B(p) = 0$.

A vonal-végpontokat megőrző vékonyító algoritmusok többféle végpont-kritériumot alkalmaznak. Az általunk javasolt eljárásnál az alábbi alkalmazzuk:

2.1. Definíció. [5] A $(\mathbb{Z}^2, 8, 4, \mathbf{B})$ képen a $p \in \mathbf{B}$ fekete pont akkor és csak akkor végpont, ha $B(p) = 1$, vagy 2.

Legyen $p \in \mathbf{B}$ a $(\mathbb{Z}^2, 8, 4, \mathbf{B})$ kép egy fekete pontja. Jelölje $C(p)$ a $(\mathbb{Z}^2, 8, 4, \mathbf{B} \cap N_8^*(p))$ (csak a p valódi fekete 8-szomszédjait tartalmazó) kép fekete komponenseinek számát. $C(p)$ egyszerűen meghatározható a Hilditch-féle keresztszám [6] segítségével:

$$X_H(p) = \sum_{i=1}^4 b_i,$$

ahol

$$b_i = \begin{cases} 1 & \text{ha } p_{2i} = 0 \text{ és } (p_{(2i+1) \bmod 8} = 1 \text{ vagy } p_{(2i+2) \bmod 8} = 1) \\ 0 & \text{különben} \end{cases},$$

ahol az $N_8(p)$ halmaz pontjait (mint Boole-változókat) az 1. ábra szerint indexeljük. Könnyen belátható, hogy tetszőleges p határpontra $C(p) = X_H(p)$.

| | | |
|-------|-------|-------|
| p_1 | p_2 | p_3 |
| p_8 | p | p_4 |
| p_7 | p_6 | p_5 |

1. ábra. $N_8(p)$ pontjainak indexelése. Megjegyezzük, hogy a 2-dimenziós merőleges képrácsot (\mathbb{Z}^2) itt és a további ábrákon is a vele duális négyzetmozaikkal [11] reprezentáljuk.

2.2. Definíció. A $(\mathbb{Z}^2, 8, 4, \mathbf{B})$ képen a $p \in \mathbf{B}$ egyszerű pont (*simple point*), ha p törlése nem változtatja meg a kép topológiáját (vagyis nem szakít szét objektumot, nem töröl teljesen egy objektumot, nem hoz létre új üreget és nem olvaszt össze üregeket sem egymással, sem pedig a háttérrel) [8].

A 2-dimenziós $(8, 4)$ képeken az egyszerűség lokális tulajdonság, mivel egy képen a p pont egyszerű volta eldönthető $N_8(p)$ ismeretében. Az egyszerűségre adott számos kritérium közül az alábbiit alkalmazzuk:

2.1. TÉTEL. [8] A $(\mathbb{Z}^2, 8, 4, \mathbf{B})$ képen a $p \in \mathbf{B}$ fekete pont akkor és csakis akkor egyszerű, ha p határpont, nem izolált pont és $C(p) = 1$.

3. Kritikus párok és tulajdonságaik

A p és q két, egymással 4-szomszédos pontokra vezessük be az $N_8(p, q) = N_8(p) \cup N_8(q)$, valamint a $N_8^*(p, q) = N_8(p, q) \setminus \{p, q\}$ jelöléseket. A továbbiakban jelölje $C(p, q)$ a fekete 8-komponensek számát $N_8^*(p, q)$ -ban. Ezenkívül a $p_1 - p_8$ jelöléseket fogjuk alkalmazni a p pont 3×3 -as környezetén belül levő pontokra, az 1. ábrán látható módon.

3.1. *Definíció.* Legyen p egy kép tetszőleges objektumpontja. A $\Gamma_i(p)$ függvényt ($i = 1, 2, 3, 4$) a következőképpen definiáljuk:

$$\Gamma_i(p) = p_{2i} \wedge (p_{(2i-2) \bmod 8} \vee p_{(2i-1)}) \wedge (p_{(2i+1) \bmod 8} \vee p_{(2i+2) \bmod 8})$$

3.2. *Definíció.* Legyenek p és q egymással 4-szomszédos határpontok egy képen. A $\{p, q\}$ halmazt kritikus párnak nevezzük, ha az alábbi feltételek teljesülnek:

- $C(p) = C(q) = 1$,
- $3 \leq B(p), B(q) \leq 6$,
- $C(p, q) = 2$.

Egy kritikus pár lehet vízszintes vagy függőleges aszerint, hogy elemei egy sorban vagy egy oszlopban helyezkednek el.

Ha egy $\{p, q\}$ kritikus párban $q \in \{p_2, p_8\}$, akkor azt mondjuk, hogy q a kritikus párban a kisebb indexű, p pedig a nagyobb indexű elem, $q \in \{p_4, p_6\}$ esetben pedig értelemszerűen ennek fordítottja érvényes.

3.1. TÉTEL. Legyenek p és q olyan egymással 4-szomszédos határpontok egy képen, melyekre érvényes, hogy:

- I. $C(p) = C(q) = 1$,
- II. $3 \leq B(p), B(q) \leq 6$.

Ekkor az alábbi állítások egymással ekvivalensek:

- (a) $\Gamma_i(p) = 1$.
- (b) $\{p, q\}$ kritikus pár.
- (c) Bármelyik $x \in \{p, q\}$ pont törlése után a megmaradó $y \in \{p, q\} \setminus \{x\}$ objektumpontra teljesül, hogy $C(y) = 2$.

Bizonyítás.

Elegendő $q = p_4$ ($i = 2$) esetet vizsgálni, mert a többi lehetséges szituációban a szimmetria miatt más jelölésekkel ugyan, de azonos módon történhet a bizonyítás. Három összefüggést bizonyítunk, melyek alapján következik a tételben megfogalmazott ekvivalencia.

(a) \Rightarrow (b). Tegyük fel először, hogy $\Gamma_2(p) = (p_2 \vee p_3) \wedge (p_5 \vee p_6) = 1$. Azt kell belátnunk, hogy $C(p, q) = 2$. Mivel $C(p) = C(q) = 1$, így a $C(p, q)$ érték megegyezik a $p_2, q_2, q_4, q_6, p_6, p_8, p_2$ sorozatban előforduló 1-0 átmenetek számával. A feltevésünk alapján $p_2 \vee q_2 = 1$ és $p_6 \vee q_6 = 1$ adódik, továbbá mivel p és q határpontok, ezért $\neg p_2 \vee \neg p_6 \vee \neg p_8 = 1$, és $\neg q_2 \vee \neg q_4 \vee \neg q_6 = 1$. Az előbbi négy összefüggésből adódik, hogy a p_2, q_2, q_4, q_6 sorozat és a q_6, p_6, p_8, p_2 sorozat is tartalmaz egy-egy 1-0 átmenetet. Továbbá 2-nél több 1-0 átmenetet nem tartalmazhat a $p_2, q_2, q_4, q_6, p_6, p_8, p_2$ sorozat, mivel akkor sérülne a $C(p) = C(q) = 1$ tulajdonság. Az előbbiekből tehát következik, hogy $C(p, q) = 2$, vagyis $\{p, q\}$ kritikus pár.

(b) \Rightarrow (a). Indirekt módon igazoljuk. Induljunk ki abból a feltevésből, hogy $\{p, q\}$ olyan kritikus pár, melyre $p_2 = p_3 = q_2 = 0$. Ekkor szükségszerűen $p_6 = p_5 = q_6 = 1$ is fennáll, különben nem teljesülne a $C(p) = C(q) = 1$ vagy a $3 \leq B(p), 3 \leq B(q)$ kikötés. Viszont ilyenkor a q_4 és p_8 értékétől függetlenül a $p_2, q_2, q_4, q_6, p_6, p_8, p_2$ sorozat csak egy darab 1-0 átmenetet fog tartalmazni, ami azt jelentené, hogy $C(p, q) = 1$, vagyis $\{p, q\}$ nem kritikus pár. Ugyanerre az ellentmondásra jutunk a $p_6 = q_6 = 0$ feltevésből is. Tehát $\Gamma_2(p) = (p_2 \vee p_3) \wedge (p_5 \vee p_6) = 1$.

(a) \Leftrightarrow (c). Tegyük fel, hogy a $\{p, q\}$ halmaz kritikus pár. $C(p) = C(q) = X_H(p) = X_H(q) = 1$, így a Hilditch-féle keresztszám definícióját alkalmazva megállapítható, hogy a $q = p_4$ törlése után a $C(p) = X_H(p) = 2$, ill. $p = q_8$ törlése után a $C(q) = X_H(q) = 2$ egyenlőség pontosan akkor áll fenn, ha $p_5 = 1$ vagy $p_6 = 1$, azaz $p_5 \vee p_6 = 1$, és $q_1 = p_2 = 1$ vagy $q_2 = p_3 = 1$, azaz $p_2 \vee p_3 = 1$. Összevonva tehát a két feltételt, (c) pontosan akkor teljesül, ha

$$\Gamma_2(p) = (p_2 \vee p_3) \wedge (p_5 \vee p_6) = 1$$

egyenlőség fennáll. □

Most tetszőleges p objektumpontra bevezetünk három kifejezést, melyeket felhasználunk a későbbi definíciókban. A harmadik képletnél feltételezzük, hogy $p_{j+8} = p_j$ ($1 \leq j \leq 8$):

$$\begin{aligned} \alpha(p) &= (\neg p_4 \wedge p_8 \wedge ((p_1 \wedge p_7) \vee (p_2 \wedge p_6))) \vee (\neg p_6 \wedge p_2 \wedge ((p_1 \wedge p_3) \vee (p_4 \wedge p_8))), \\ \beta(p) &= (\neg p_2 \wedge p_6 \wedge ((p_5 \wedge p_7) \vee (p_4 \wedge p_8))) \vee (\neg p_8 \wedge p_4 \wedge ((p_3 \wedge p_5) \vee (p_2 \wedge p_6))), \\ \gamma(p) &= \bigvee_{i=1}^4 (p_{2i} \wedge p_{2i+2} \wedge \neg p_{2i+1} \wedge \neg p_{2i+4} \wedge \neg p_{2i+5} \wedge \neg p_{2i+6}). \end{aligned}$$

3.3. Definíció. Egy kép p objektumpontját rendre α -pontnak, β -pontnak, γ -pontnak nevezzük, ha $\alpha(p) = 1$, $\neg\alpha(p) \wedge \beta(p) = 1$, $\gamma(p) = 1$.

Az α -, β -, γ -pontokat a 2. ábrán szemléltetjük.

A kritikus párokat kritikus α -párnak, β -párnak, ill. γ -párnak fogjuk nevezni, ha mindkét elemük α -, β -, ill. γ -pont, továbbá $\alpha\beta$ -, $\alpha\gamma$ -, ill. $\beta\gamma$ -párokról fogunk beszélni, ha bennük éppen a megjelölt 2-2 fajta pont szerepel.

3.1. SEGÉDTÉTEL. Egyetlen objektumpont sem szerepelhet egynél több kritikus α -párban.

Bizonyítás. Tegyük fel, hogy $\{p, q\}$ kritikus α -pár, ahol $q = p_4$ (3. ábra). A 3.1. tétel szerint:

$$(p_2 \vee q_2) \wedge (p_6 \vee q_6) = 1. \quad (1)$$

Mivel $p_4 = q = 1$, ezért a 3.3. definíció szerint:

$$\alpha(p) = \neg p_6 \wedge p_2 \wedge ((p_1 \wedge p_3) \vee (p_8 \wedge p_4)) = \neg p_6 \wedge p_2 \wedge ((p_1 \wedge q_2) \vee p_8) = 1. \quad (2)$$

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--|-----|-----|-----|---|-----|---|---|-----|---|---|---|---|---|---|-----|---|---|---|---|--|-----|---|---|-----|-----|-----|---|---|---|---|---|---|---|---|-----|---|---|---|---|
| <table><tr><td>1</td><td>.</td><td>.</td></tr><tr><td>1</td><td>p</td><td>0</td></tr><tr><td>1</td><td>.</td><td>.</td></tr></table> | 1 | . | . | 1 | p | 0 | 1 | . | . | <table><tr><td>.</td><td>1</td><td>.</td></tr><tr><td>1</td><td>p</td><td>0</td></tr><tr><td>.</td><td>1</td><td>.</td></tr></table> | . | 1 | . | 1 | p | 0 | . | 1 | . | <table><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>.</td><td>p</td><td>.</td></tr><tr><td>.</td><td>0</td><td>.</td></tr></table> | 1 | 1 | 1 | . | p | . | . | 0 | . | <table><tr><td>.</td><td>1</td><td>.</td></tr><tr><td>1</td><td>p</td><td>1</td></tr><tr><td>.</td><td>0</td><td>.</td></tr></table> | . | 1 | . | 1 | p | 1 | . | 0 | . |
| 1 | . | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | p | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | . | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | 1 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | p | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | 1 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | p | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | 0 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | 1 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | p | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | 0 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (a) | (b) | (c) | (d) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table><tr><td>x</td><td>y</td><td>1</td></tr><tr><td>0</td><td>p</td><td>1</td></tr><tr><td>.</td><td>z</td><td>1</td></tr></table> | x | y | 1 | 0 | p | 1 | . | z | 1 | <table><tr><td>.</td><td>1</td><td>.</td></tr><tr><td>0</td><td>p</td><td>1</td></tr><tr><td>.</td><td>1</td><td>.</td></tr></table> | . | 1 | . | 0 | p | 1 | . | 1 | . | <table><tr><td>x</td><td>0</td><td>.</td></tr><tr><td>y</td><td>p</td><td>z</td></tr><tr><td>1</td><td>1</td><td>1</td></tr></table> | x | 0 | . | y | p | z | 1 | 1 | 1 | <table><tr><td>.</td><td>0</td><td>.</td></tr><tr><td>1</td><td>p</td><td>1</td></tr><tr><td>.</td><td>1</td><td>.</td></tr></table> | . | 0 | . | 1 | p | 1 | . | 1 | . |
| x | y | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | p | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | z | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | 1 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | p | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | 1 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| x | 0 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| y | p | z | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | 0 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | p | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | 1 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (e) | (f) | (g) | (h) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table><tr><td>.</td><td>1</td><td>0</td></tr><tr><td>0</td><td>p</td><td>1</td></tr><tr><td>0</td><td>0</td><td>.</td></tr></table> | . | 1 | 0 | 0 | p | 1 | 0 | 0 | . | <table><tr><td>0</td><td>0</td><td>.</td></tr><tr><td>0</td><td>p</td><td>1</td></tr><tr><td>.</td><td>1</td><td>0</td></tr></table> | 0 | 0 | . | 0 | p | 1 | . | 1 | 0 | <table><tr><td>.</td><td>0</td><td>0</td></tr><tr><td>1</td><td>p</td><td>0</td></tr><tr><td>0</td><td>1</td><td>.</td></tr></table> | . | 0 | 0 | 1 | p | 0 | 0 | 1 | . | <table><tr><td>0</td><td>1</td><td>.</td></tr><tr><td>1</td><td>p</td><td>0</td></tr><tr><td>.</td><td>0</td><td>0</td></tr></table> | 0 | 1 | . | 1 | p | 0 | . | 0 | 0 |
| . | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | p | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | p | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | p | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | p | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| . | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (i) | (j) | (k) | (l) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

2. ábra. Az (a)-(d) ábrákon az α -pontok, az (e)-(h) ábrákon a β -pontok, az (i)-(l) ábrákon pedig a γ -pontok konfigurációi láthatók. Az (e) és (g) konfigurációkban $x \wedge y = 0$ vagy $z = 1$ kell, hogy teljesüljön. A "." helyén állhat akár 0, akár 1.

(2) alapján $p_6=0$, ezután (1)-ből $q_6 = 1$ adódik, és mivel a 3.3. definíció szerint $\alpha(q) = 1$, következik, hogy

$$\neg q_4 \wedge q_2 = 1. \quad (3)$$

Legyen $r = p_2$. Könnyen látható, hogy $\alpha(r) = 1$ csak akkor teljesülhet, ha $\neg r_4 \vee \neg r_6 = 1$. De mivel $r_6 = p = 1$ és (3) szerint $r_4 = q_2 = 1$, ezért biztos, hogy p_2 nem lehet α -pont.

$r = p_8$ -re hasonlóan észrevehető, hogy ha p_8 α -pont, akkor $r_6 = p_7 = 0$. Viszont mivel $p_6 = 0$, így ekkor $\Gamma_2(r) = 0$ lenne, ezért a 3.1. tétel szerint $\{p, p_8\}$ nem alkothat kritikus α -párt. Ugyanígy belátható, hogy ha $r = q_2$ és r α -pont, akkor $\Gamma_3(r) = 0$, ezért $\{q, q_2\}$ nem alkothat kritikus α -párt.

Végül legyen $r = q_6$. Mivel fentebb már láttuk, hogy $q_4 = r_3 = 0$ és $p_6 = r_8 = 0$, ezért $\alpha(r) = 0$, így q_6 sem lehet α -pont.

Ezzel végigvizsgáltuk az összes lehetséges p -vel vagy q -val szomszédos objektumpontot, így ha $\{p, q\}$ vízszintes helyzetű α -pár, akkor valóban teljesül az állításunk. Függőleges α -pár esetén a segédétel hasonló módon bizonyítható. \square

3.2. SEGÉDTÉTEL. Egyetlen objektumpont sem szerepelhet egynél több kritikus β -párban.

Bizonyítás. Vegyük észre, hogy ha a β -pontok környezetét 180 fokkal elforgatjuk, akkor éppen α -pontokat kapunk. Ezért a bizonyítás (eltérő jelölésekkel) ugyanúgy végezhető, mint az előző állítás esetén. \square

| | | | |
|-------|-------|-------|-------|
| p_1 | p_2 | q_2 | q_3 |
| p_8 | p | q | q_4 |
| p_7 | p_6 | q_6 | q_5 |

3. ábra. A $\{p, q\}$ halmaz szomszédsága $q = p_4$ esetben.

3.3. SEGÉDTÉTEL. Legyen $\{p, q\}$ egy kritikus β -pár, $\{r, s\}$ pedig egy kritikus α -pár, ahol q nagyobb indexű p -nél, s pedig nagyobb indexű r -nél. Ha p -nek 4-szomszédja r , akkor a $\{p, r\}, \{q, s\}$ halmazok $\alpha\beta$ -párok.

Bizonyítás. Legyen $\{p, q\}$ kritikus β -pár. Ismét csak a 3. ábrán látható szituációt elemezzük, a többi eset a szimmetriák miatt erre visszavezethető. Akkor ismét fennáll az (1) összefüggés, továbbá a 3.3. definíció szerint:

$$\beta(q) = \neg q_2 \wedge q_6 \wedge ((q_5 \wedge q_7) \vee q_4) = 1.$$

Ebből $q_6 = 1$ adódik. Ha $r = p_2$ lenne, akkor s szükségszerűen csak kisebb indexű lehetne r -nél, de ez sértené a tételben szereplő kikötést. Tehát csak $r = p_6$ lehet. $r_4 = q_6 = 1$ miatt a következő összefüggés teljesül:

$$\alpha(r) = \neg r_6 \wedge r_2 \wedge ((r_1 \wedge r_3) \vee (r_4 \wedge r_8)) = \neg r_6 \wedge ((p_5 \wedge p_7) \vee p_8) = 1. \quad (4)$$

Ha $p_8 = 1$ lenne, akkor $p_2 = 0$ kellene, hogy teljesüljön, különben $\beta(p) = 0$ lenne, azaz p nem lenne β -pont, ami ellentmond az állításban megfogalmazott feltevésnek. Tehát $p_8 = 0$, így ebből és a (4) összefüggésből következik, hogy $p_5 \wedge p_7 = 1$. Az előbbiek alapján $s = r_8$ esetén $s_4 = r = 1$, és $s_2 = 0$ miatt $\alpha(s) = 0$ lenne, ezért csak $s = r_4 = q_6$ lehet, és könnyen ellenőrizhető, hogy ekkor valóban teljesül az $\alpha(s) = 1$ feltétel. Tehát a $\{p, r\}, \{q, s\}$ halmazok valóban $\alpha\beta$ -párok. \square

3.4. SEGÉDTÉTEL. Ha p kritikus pár eleme, akkor p csak α -, β -, vagy γ -pont lehet.

Bizonyítás. Csak arra az esetre adunk bizonyítást, amikor $\{p, p_4\}$ alkot kritikus párt, a többi szituáció ehhez hasonlóan vizsgálható. Ha $\{p, p_4\}$ kritikus pár, akkor

$$\Gamma_2(p) = (p_2 \vee p_3) \wedge (p_5 \vee p_6) = 1. \quad (5)$$

Tegyük fel, hogy p nem α -pont, és nem β -pont. Azt kell igazolnunk, hogy ekkor p γ -pont. A De Morgan-azonosságok alkalmazásával adódik, hogy

$$\begin{aligned} \neg\alpha(p) \wedge \neg\beta(p) &= \\ &= (p_4 \vee \neg p_8 \vee ((\neg p_1 \vee \neg p_7) \wedge (\neg p_2 \vee \neg p_6))) \wedge \\ &\quad (p_6 \vee \neg p_2 \vee ((\neg p_1 \vee \neg p_3) \wedge (\neg p_4 \vee \neg p_8))) \wedge \\ &\quad (p_2 \vee \neg p_6 \vee ((\neg p_5 \vee \neg p_7) \wedge (\neg p_4 \vee \neg p_8))) \wedge \\ &\quad (p_8 \vee \neg p_4 \vee ((\neg p_3 \vee \neg p_5) \wedge (\neg p_2 \vee \neg p_6))) = 1. \end{aligned} \quad (6)$$

Mivel $p_4 = 1$, ezért $(p_8 \vee ((\neg p_3 \vee \neg p_5) \vee (\neg p_2 \vee \neg p_6))) = 1$. Ha $p_8 = 1$, akkor $p_2 \vee p_6 = 1$, különben $C(p) > 1$ lenne. $p_2 = p_6 = 1$ esetben p nem lenne határpont, így e két pont közül pontosan az egyiknek kell objektumpontnak lennie. Ekkor azonban a (6)-ben szereplő konjunkcióban a középső 2 tag valamelyikének 0 lesz az értéke, így viszont (6) nem teljesülhetne. Tehát $p_8 = 0$, és

$$(\neg p_3 \vee \neg p_5) \wedge (\neg p_2 \vee \neg p_6) = 1. \quad (7)$$

(5)-ből és (7)-ből következik, hogy $(p_2 \wedge \neg p_3 \wedge p_5 \wedge \neg p_6) \vee (\neg p_2 \wedge p_3 \wedge \neg p_5 \wedge p_6) = 1$.

$(p_2 \wedge \neg p_3 \wedge p_5 \wedge \neg p_6) = 1$ esetben $p_7 = 0$, különben $C(p) > 1$ lenne. Ekkor tehát $(p_2 \wedge p_4 \wedge \neg p_3 \wedge \neg p_6 \wedge \neg p_7 \wedge \neg p_8) = 1$.

$(\neg p_2 \wedge p_3 \wedge \neg p_5 \wedge p_6) = 1$ esetben $p_1 = 0$, különben $C(p) > 1$ állna fenn. Így ekkor $(p_4 \wedge p_6 \wedge \neg p_5 \wedge \neg p_8 \wedge \neg p_1 \wedge \neg p_2) = 1$.

Mindkét előbbi esetben megállapítható, hogy $\gamma(p) = 1$, tehát a 3.3. definíció szerint p γ -pont. \square

3.4. Definíció. p -t biztonságos γ -pontnak nevezzük, ha γ -pont, és

$$\gamma^*(p) = (\neg p_4 \wedge \neg p_5 \wedge \neg p_6 \wedge \neg p_1 \wedge p_8 \wedge p_2) \vee (\neg p_6 \wedge \neg p_7 \wedge \neg p_8 \wedge \neg p_3 \wedge p_2 \wedge p_4) = 1.$$

A biztonságos γ -pontokat a 4. ábra szemlélteti.

| | | |
|---|-----|---|
| 0 | 1 | . |
| 1 | p | 0 |
| . | 0 | 0 |

(a)

| | | |
|---|-----|---|
| . | 1 | 0 |
| 0 | p | 1 |
| 0 | 0 | . |

(b)

4. ábra. Konfigurációk, ahol p biztonságos γ -pont.

3.5. SEGÉDTÉTEL. Kritikus γ -párban pontosan az egyik pont biztonságos γ -pont.

Bizonyítás. Ha $\{p, q\}$ kritikus párban p biztonságos γ -pont, akkor $\gamma^*(p) = 1$. Egyszerűen belátható, hogy $\gamma(q) = 1$ csak úgy teljesülhet, ha

$$(\neg q_1 \wedge \neg q_2 \wedge \neg q_8 \wedge \neg q_5 \wedge q_4 \wedge q_6) \vee (\neg q_2 \wedge \neg q_3 \wedge \neg q_4 \wedge \neg q_7 \wedge q_6 \wedge q_8) = 1,$$

azaz ha $\gamma(q)^* = 0$. A $\gamma^*(p) = 0$ esetben is hasonlóan következtethetünk $\gamma^*(q) = 1$ -re. \square

4. A javasolt algoritmus

Az ismertetésre kerülő algoritmusunkban a bemeneti képet tároló P képmátrix mellett egy azzal megegyező méretű M segédmátrixot is bevezetünk abból a célból, hogy tároljuk a törölt vagy a törlésre esélyes képpontokat. A $p \in P$ és az $m \in M$ egymással megegyező indexű pontokra bevezetjük a következő jelöléseket:

$$\neg N_4^*(p) = \{\neg p_2, \neg p_4, \neg p_6, \neg p_8\},$$

$$N_4^*(p) \wedge N_4^*(m) = \{p_2 \wedge m_2, p_4 \wedge m_4, p_6 \wedge m_6, p_8 \wedge m_8\}.$$

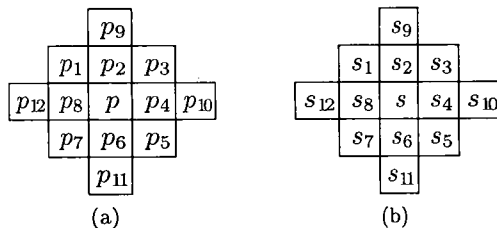
A H képpontokat tartalmazó halmazra jelölje $|H|$ a H -ban levő 1 értékű pontok számát.

Az algoritmusunk hatékonysága érdekében a két képmátrix (P és M) mellett két halmazt, R -t és Q -t is használunk, amelyekbe gyűjthetjük a számunkra érdekes képpontokat. Mivel nincs szükségünk bonyolultabb halmazműveletekre, így az R és a Q segédhalmazokat egy-egy láncolt listában is tárolhatjuk.

A javasolt algoritmus az iteratív objektum redukció előtt inicializálja az R halmazt és az M segédmátrixot. Az R halmazba bekerülnek a P -ben tárolt vékonyítandó kép objektumpontjai, az M segédmátrixnak pedig minden elemét 0-ra állítjuk.

A vékonyító eljárás egy iterációs lépése két menetből áll. Az első menetben megvizsgáljuk az R halmazban tárolt objektumpontokat. Ha találunk R -ben olyan nem izolált objektumpontot, amely megjelölt (1-re állított) M -ben, akkor azt már ebben a menetben törölhetjük. Ha pedig a vizsgált képpont határpont az aktuális P képen, de M -ben nincs megjelölve, akkor berakjuk őt a Q halmazba. Az első menet végén M minden elemét 0-ra állítjuk.

Az iterációs lépések második menetében sorra megvizsgáljuk a Q halmaz pontjait. Bevezetjük a p képpontra a 8-szomszédságnál bővebb, (p -n kívül) 12 pontot tartalmazó környezetet, ahol a p_i ($9 \leq i \leq 12$) pontok az 5(a). ábra szerint helyezkednek el. A második menetben a vizsgált pontok ezen bővebb környezetét vesszük figyelembe a P és az M mátrixokban. Szükségünk van még az s (Boole) változóra és a 12-elemű $S = (s_1, s_2, \dots, s_{12})$ bitvektorra, ahol $s = p \vee m$ és $s_i = p_i \vee m_i$ ($1 \leq i \leq 12$), vagyis s a vizsgált p képpont és a vele megegyező indexű M -beli pont



5. ábra. A p pont 12 elemű környezete P -ben, valamint az s pont és vizsgálandó S környezete.

összevagyoltja, S pedig p és m 12-elemű környezeteinek összevagyolásával kapható. S segítségével visszanyerhetjük p -nek az adott iteráció elején fennálló környezetét és eldönthetjük a vizsgált p képpont törölhetőségét. Az algoritmus pszeudo-kódjában szereplő $C(s)$, $B(s)$, $\Gamma_j(s)$ ($1 \leq j \leq 4$) jelölések az S -ben tárolt környezetnek (5(b). ábra) megfelelően értelmezendők.

A fenti jelölések segítségével az alábbiakban megadjuk a javasolt algoritmusunk pszeudo-kódját:

Input: a kiindulási bináris képet tároló P bitmátrix.

Output: a P képmátrix, amelyben a vázat tartalmazó bináris kép alakul ki.

Segédátrolók:

- a Q és az R halmazok,
- az M bitmátrix (melynek mérete megegyezik a P képmátrixéval),
- az $S = (s_1, s_2, \dots, s_{12})$ bitvektor.

Inicializálás:

$R := \emptyset$

$Q := \emptyset$

for minden $p \in P$ elemre do

if $p = 1$ then

$R := R \cup \{p\}$

for minden $m \in M$ elemre do

$m := 0$

repeat

1. menet:

for minden $p \in R$ képpontra do

Legyen m a p -vel megegyező indexű pont M -ben.

if $B(p) > 0$ and $m = 1$ then

$p := 0$

$R := R \setminus \{p\}$

else if $0 < |N_4^*(p) \cap \neg N_4^*(m)| < 4$ then

$Q := Q \cup \{p\}$

$m := 0$

2. menet:

for minden $p \in Q$ képpontra do

$Q := Q \setminus \{p\}$

Legyen m a p -vel megegyező indexű pont M -ben.

$s := p \vee m$

if $s = 1$ then

for $i = 1, 2, \dots, 12$ do

$s_i := p_i \vee m_i$

if $C(p) = 1$ and $C(s) = 1$ and $3 \leq B(s) \leq 6$ and
 $\neg\gamma(s) \vee \gamma^*(s) = 1$ then
 if teljesül az $F_1(s) - F_6(s)$ feltételek valamelyike then
 $p := 0$
 $R := R \setminus \{p\}$
 $m := 1$
 for $i = 1, 2, 3, 4$ do
 $m_{2i} := m_{2i} \wedge \neg\Gamma_i(s)$
 else
 $m := \neg\gamma(s) \wedge \left(\bigwedge_{i=1}^4 (p_{2i} \vee \neg m_{2i} \vee \neg\Gamma_i(s)) \right)$
 until nem történt változás

Az $F_i(s)$ ($i = 1, 2, \dots, 6$) feltételeket a következő formulák és a 6. ábra maszkjai segítségével adjuk meg:

$F_1(s)$: S illeszkedik a 6(a)-6(b) ábrákon feltüntetett maszkok egyikére, ahol

$$\neg(x_1 \wedge x_2 \wedge \neg x_3) = 1.$$

$F_2(s)$: S illeszkedik a 6(c)-6(d) ábrákon feltüntetett maszkok egyikére, ahol

$$\neg(x_1 \wedge x_2 \wedge \neg x_5) \wedge \neg(x_1 \wedge \neg x_2 \wedge x_5) \wedge \neg(x_2 \wedge \neg x_3) \wedge \\ \wedge \neg(\neg x_3 \wedge x_4 \wedge x_6) \wedge \neg(x_3 \wedge x_4 \wedge \neg x_6) = 1.$$

$F_3(s)$: S illeszkedik a 6(e) ábrán szereplő maszkra, ahol

$$\neg(\neg x_1 \wedge x_2 \wedge x_4) \wedge \neg(\neg x_1 \wedge x_3 \wedge x_5) = 1.$$

$F_4(s)$: S illeszkedik a 6(f)-6(g) ábrákon levő maszkok valamelyikére, ahol

$$\neg(\neg x_1 \wedge x_2 \wedge x_4) \wedge \neg(\neg x_1 \wedge x_3 \wedge x_5) \wedge \neg(x_1 \wedge x_3 \wedge \neg x_5) = 1.$$

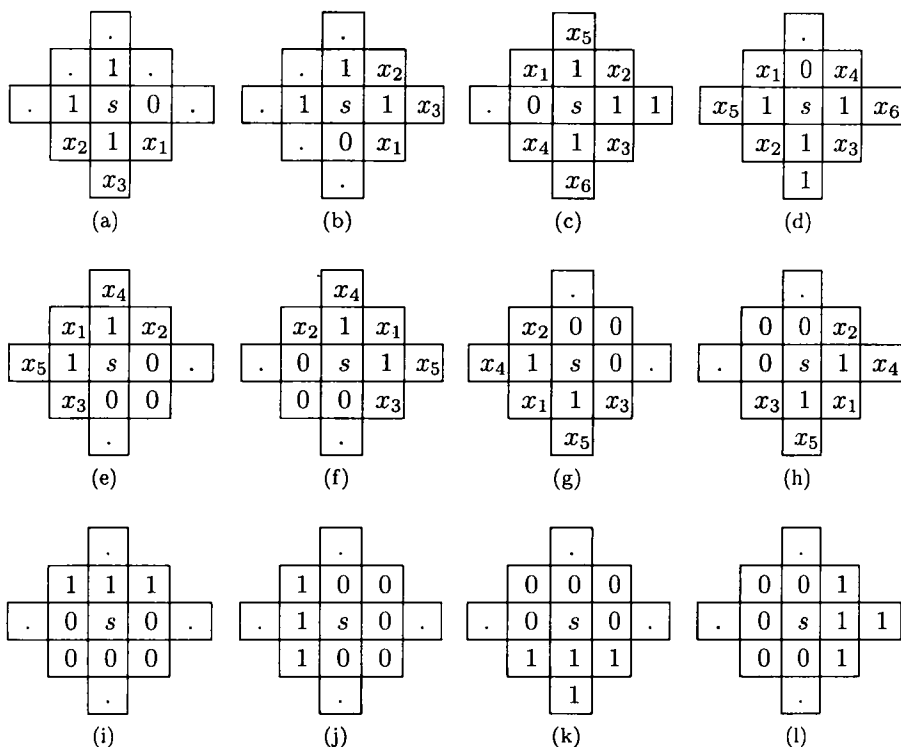
$F_5(s)$: S illeszkedik a 6(h) ábrán feltüntetett maszkra, ahol

$$\neg(\neg x_1 \wedge x_2 \wedge x_4) \wedge \neg(x_1 \wedge x_2 \wedge \neg x_4) \wedge \\ \wedge \neg(\neg x_1 \wedge x_3 \wedge x_5) \wedge \neg(x_1 \wedge x_3 \wedge \neg x_5) = 1,$$

és

$$(x_1, x_2, x_3, x_4, x_5) \neq (1, 0, 0, 0, 0).$$

$F_6(s)$: S illeszkedik a 6(i)-6(l) ábrákon feltüntetett maszkok egyikére.



6. ábra. Az $F_1(s) - F_6(s)$ feltételekhez tartozó törlőmaszkok, ahol $s = p \vee m = 1$. Ha a maszk változót tartalmaz, akkor az illeszkedéshez a feltételekben megadott logikai kifejezéseknek is teljesülniük kell. A "." ("don't care") maszkpozíciók fekete és fehér pontokra egyaránt illeszkednek.

5. Az algoritmus tulajdonságainak igazolása

Az alábbiakban igazolni fogjuk, hogy az ismertett algoritmusunk bejárás-független, topológia-megőrző, és bizonyos értelemben maximálisan vékonyít (utóbbinál az egyértelműség végett előbb definiálni fogjuk, hogy mit értünk maximálisan vékonyító algoritmuson). Előzetesen bevezetjük az ideális pontok fogalmát, melyekkel kapcsolatosan felírhatók bizonyos összefüggések a kritikus párokra, ill. az algoritmusunkra vonatkozóan. Ezeket felhasználjuk az említett tulajdonságok bizonyításánál.

5.1. Definíció. A p objektumpontot megmaradó sarokpontnak nevezzük, ha p illeszkedik a 7. ábrán levő maszkra.

5.2. Definíció. Egy p képpontot ideális pontnak nevezünk, ha az teljesíti a következő feltételeket:

- $3 \leq B(p) \leq 6$,
- $C(p) = 1$,
- teljesül az alábbi $G_1(p) - G_4(p)$ kritériumok egyike.

$G_1(p)$: p nem eleme kritikus párnak, és nem megmaradó sarokpont.

$G_2(p)$: p olyan α -pont, amely nem egy kritikus α -pár kisebb indexű tagja.

$G_3(p)$: p olyan β -pont, amelyre teljesülnek az alábbi feltételek:

- i)* ha p egy β -párnak nagyobb indexű eleme, akkor ezen párba tartozó másik pont egy $\alpha\beta$ -párnak is eleme.
- ii)* ha p egy $\alpha\beta$ -párnak is eleme, akkor ezen párnak α -pontja egyben egy α -párnak kisebb indexű eleme.

$G_4(p)$: p olyan biztonságos γ -pont, amely sem $\alpha\gamma$ -párnak, sem pedig $\beta\gamma$ -párnak nem eleme.

| | | | |
|---|-----|---|---|
| 0 | 0 | 0 | 0 |
| 0 | p | 1 | 0 |
| 0 | 1 | 1 | 0 |
| 0 | 0 | 0 | 0 |

7. ábra. A megmaradó p sarokpont környezete. (p egy izolált 2×2 -es "négyzet" bal felső eleme.)

5.1. SEGÉDTÉTEL. *Nincs olyan kritikus pár egy képen, melynek mindkét eleme ideális pont.*

Bizonyítás. Legyen $\{p, q\}$ kritikus pár, melyben p kisebb indexű q -nál. Ha $\{p, q\}$ α -pár, akkor nyilván p nem ideális pont, mivel $G_2(p)$ nem teljesül.

Tegyük fel, hogy $\{p, q\}$ β -pár, és teljesül $G_3(p)$. A 3.2. segédtétel alapján q nem lehet egy további kritikus β -párnak is az eleme, így $G_3(q)$ *i)* részfeltétele csak úgy teljesülhet, ha p eleme egy $\{p, r\}$ kritikus $\alpha\beta$ -párnak is. $G_3(p)$ szerint viszont ekkor kell, hogy legyen olyan $\{r, s\}$ kritikus α -pár, amelyben r kisebb indexű s -nél. A 3.3. segédtételből pedig így az következne, hogy $\{q, s\}$ is egy kritikus $\alpha\beta$ -pár, emiatt nem teljesülne $G_3(q)$ *ii)* részfeltétele, így q nem lehet ideális pont. Ha pedig abból indulunk ki, hogy $G_3(q)$ teljesül, akkor az előbbi esethez hasonlóan, a 3.3. segédtétel felhasználásával következtethetünk arra, hogy p nem ideális pont.

Most tegyük fel, hogy $\{p, q\}$ $\alpha\beta$ -pár. Ha p az α -pont, és $G_2(p)$ teljesül, akkor $G_3(q)$ nyilván nem teljesülhet a *ii*) feltétel megsértése miatt. Ugyanígy, ha q az α -pont, és $G_2(q)$ teljesül, akkor $G_3(p)$ nem teljesülhet. Tehát ekkor sem lehet mindkét pont ideális.

Végül ha p biztonságos γ -pont, és q nem γ -pont, akkor a 3.4. segédttétel szerint $G_4(p)$ nem teljesül, ha pedig q γ -pont, akkor a 3.5. segédttétel szerint q nem biztonságos γ -pont, így q nem lehet ideális pont. Ehhez hasonlóan belátható az is, hogy ha q biztonságos γ -pont, akkor p nem ideális pont. \square

5.2. SEGÉDTÉTEL. *Ha p α -pont vagy β -pont, de nem ideális pont, akkor p eleme egy olyan $\{p, q\}$ kritikus párnak, amelyben q ideális pont.*

Bizonyítás. Tegyük fel, hogy p nem ideális pont. Ha p α -pont, akkor mivel $G_2(p)$ nem teljesül, ezért p biztosan kisebb indexű eleme egy $\{p, q\}$ kritikus α -párnak, tehát a 3.1. segédttétel szerint $G_2(q)$ teljesül, azaz q ideális pont.

Ha p β -pont, akkor mivel $G_3(p)$ nem teljesül, ezért vagy *i*), vagy *ii*) feltétel nem teljesül. Előbbi esetben p egy $\{p, q\}$ β -párnak nagyobb indexű eleme, amelyben q nem eleme $\alpha\beta$ -párnak, vagyis a 3.2. segédttétel szerint teljesül $G_3(q)$. *ii*) feltétel sérülése esetén p egy $\{p, q\}$ $\alpha\beta$ -párnak eleme, amelyben q vagy nem eleme α -párnak, vagy egy ilyen párban a nagyobb indexű elemként szerepel. Következésképpen teljesül $G_2(q)$, azaz q ezúttal is ideális pont. \square

A következő állítás igazolásához bevezetjük a $P' = P \vee M$ mátrixot

$$(P'(x, y) = P(x, y) \vee M(x, y), \text{ minden } (x, y) \text{ képpontra}).$$

A $p' \in P'$ pont indexe azonos lesz a $p \in P$ és az $m \in M$ ponttal. Mivel adott $p \in P$ esetén az algoritmusban használt S vektor elemei megegyeznek p' 12 elemű környezetének elemeivel, ezért az $s = p \vee m$ jelölést a továbbiakban p' -vel fogjuk helyettesíteni. Így az $F_1(p') - F_6(p')$ feltételek is hasonlóan értelmezendők p' környezetére, mint az algoritmus leírásában definiált $F_1(s) - F_6(s)$ feltételek.

5.1. TÉTEL. *A javasolt algoritmus akkor és csak akkor törli a p képpontot, ha valamely iterációjának elején p ideális pont.*

Bizonyítás. Ha $C(p) \neq 1$, vagy $C(p') \neq 1$, vagy $B(p') < 3$, vagy $B(p') > 6$, vagy $\neg\gamma^*(p') \wedge \gamma(p) = 1$, akkor nyilván p nem lenne ideális pont, és az algoritmus sem törölné azt. Ezért csak olyan p pontokat kell vizsgálnunk, amelyek meglátogatásakor $C(p) = 1$, $C(p') = 1$, $3 \leq B(p') \leq 6$, és $\gamma^*(p') \vee \neg\gamma(p') = 1$.

Tegyük fel először, hogy p -t törli az algoritmus valamely iteráció 2. menetében.

Ha p' α -pont, akkor az $F_1(p')$, $F_3(p')$, $F_4(p')$, $F_6(p')$ feltételek valamelyike kell, hogy teljesüljön, ugyanis az ezekben említett maszkokra történő illeszkedés esetén teljesülhet az $\alpha(p') = 1$ egyenlőség. Vegyük észre, hogy $F_1(p')$ -ben az $\neg(x_1 \wedge x_2 \wedge \neg x_3)$, míg $F_4(p')$ -ben a $\neg(x_1 \wedge x_3 \wedge \neg x_5)$ kifejezések értéke pontosan akkor 0, ha p' egy kritikus α -pár kisebb indexű eleme. Viszont ha az $F_1(p')$ -ben

feltüntetett egyenlőség teljesül, akkor nyilván az előbbi kifejezések értéke is 1, így ekkor p' csak nagyobb indexű elem lehet egy ilyen párban. Továbbá ha p' a 6(e), 6(i), 6(j) ábrákon levő maszkok valamelyikére illeszkedik, akkor a maszkon levő helyzete alapján szintén biztos, hogy az nem lesz kisebb indexű elem egy kritikus α -párban. Tehát $F_1(p')$, $F_3(p')$, $F_4(p')$, $F_6(p')$ feltételek egyikének teljesülésekor a p' α -pont ideális pont lesz, mivel teljesül rá $G_1(p')$ vagy $G_2(p')$.

Most induljunk ki abból, hogy p' β -pont. Ekkor az $F_2(p')$, $F_4(p')$, $F_5(p')$, $F_6(p')$ feltételek valamelyike kell, hogy teljesüljön, mivel csak a 6(c), 6(d), 6(f), 6(g), 6(h), 6(k), 6(l) ábrák maszkjainak egyikére illeszkedő p' -re teljesülhet $\beta(p')$. Könnyen ellenőrizhető, hogy az $F_2(p')$, $F_4(p')$, $F_5(p')$ -n belüli konjunkciókban a zárójeles tagok negációjának értéke pontosan akkor 0, ha p' -ről elmondható, hogy egy kritikus β -pár nagyobb indexű eleme, vagy egy kritikus $\alpha\beta$ -párnak eleme. Ha pedig p' környezete a 6(k), 6(l) ábrák egyikének felel meg, akkor biztos, hogy p' nem eleme kritikus párnak. Következik tehát, hogy a p' β -pontra $G_1(p')$ vagy $G_3(p')$ teljesül, azaz p' ideális pont.

Most tekintsük azt az esetet, amikor p' γ -pont. Ekkor p' a 6(e)–6(h) ábrákon levő valamely maszkra illeszkedhet, vagyis $F_3(p')$, $F_4(p')$, $F_5(p')$ feltételek egyike kell, hogy teljesüljön. Ez esetben megfigyelhető, hogy a felsorolt feltételekhez tartozó konjunkcióban a zárójeles tagok negációjának értéke pontosan akkor 0, ha p' kritikus párt alkot egy α -ponttal vagy egy β -ponttal. Az is nyilvánvaló, hogy p' biztonságos γ -pont, különben nem teljesülne a $\neg\gamma(p') \vee \gamma^*(p') = 1$ feltétel. Ebből már következik, hogy a p' γ -pontra $G_1(p')$ vagy $G_4(p')$ teljesül, vagyis p' ekkor is ideális pont.

Az előbbieket alapján az is elmondható, hogy ha az említett konjunkciókban egyik zárójeles tag negációja sem 0 értékű, akkor p' vagy nem eleme kritikus párnak, vagy biztonságos γ -pont, amely nem eleme kritikus $\alpha\gamma$ - ill. $\beta\gamma$ -párnak. Továbbá ha p' szomszédsága megfelel a 6(h) ábrának, akkor $(x_1, x_2, x_3, x_4, x_5) \neq (1, 0, 0, 0, 0)$ miatt p' nem megmaradó sarokpont (a többi maszkon pedig egyértelműen látható, hogy azok biztosan nem egy megmaradó sarokpont környezetét mutatják).

A fentiekből megállapítható tehát, hogy a törlési feltételek teljesülésekor p' biztosan ideális pont, azaz p ideális pont volt az iteráció elején. Most tegyük fel, hogy az említett törlési feltételek nem teljesülnek, de p eleme egy tetszőleges $\{p, q\}$ kritikus párnak, és p törlésre kerül a következő iteráció 1. menetében. Vegyük észre, hogy a

$$\neg\gamma(p') \wedge \left(\bigwedge_{i=1}^4 (p_{2i} \vee \neg m_{2i} \vee \neg\Gamma_i(p')) \right)$$

kifejezés értéke pontosan akkor 1, ha p' nem γ -pont és p -nek az iteráció elején nem volt olyan 4-szomszédja, amely már korábban törlésre került és kritikus párt alkotott p -vel. Ezért ha q -t p előtt látogatnánk meg, és q -t törölnénk, akkor $m = 0$ maradna, ha pedig q -t p után vizsgálnánk meg, és úgy törölnénk, akkor az $m_{2i} := m_{2i} \wedge \neg\Gamma_i(p')$ értékadások miatt lenne újból $m = 0$. Így viszont mindkét esetben ellentmondásra jutnánk abból a feltevésből, hogy p törlésre kerül,

hiszen a következő iteráció 1. menetében pontosan azon pontokat törli az algoritmus, amelyek M -beli segédértéke 1. Tehát q -t nem törölhattük az adott iterációban, vagyis a q -val megegyező indexű $q' \in P \vee M$ -re $F_1(q') - F_6(q')$ sem teljesül. Könnyen belátható, hogy ha p α -pont lett volna, akkor a 2. menetben teljesültek volna rá a törlési feltételek, de ezúttal ez nem lehetséges. Emellett a fentebb említett kifejezésben szereplő $\neg\gamma(p')$ feltétel miatt p γ -pont sem lehetett, tehát a 3.4. segédétel szerint p β -pont kellett, hogy legyen. Ha q α -pont, akkor egy kritikus α -párban kellett, hogy szerepeljen a kisebb indexű elem helyén, különben teljesült volna $F_1(q')$, $F_3(q')$, $F_4(q')$, $F_6(q')$ valamelyike. Ha q β -pont, akkor pedig kritikus $\alpha\beta$ -párnak eleme, vagy egy kritikus β -párban a nagyobb indexű elem kellett, hogy legyen, különben $F_2(q')$, $F_4(q')$, $F_5(q')$, $F_6(q')$ közül teljesült volna az egyik. Ebből következik, hogy $G_3(p)$ teljesült az iteráció elején, tehát p ebben az esetben is ideális pont volt.

Ha p nem teljesíti a törlési feltételeket, és nem eleme kritikus párnak, de törlésre kerül a következő iteráció 1. menetében, akkor ez csak úgy lehetséges, hogy p' környezete a 6(h) ábra maszkjának felelt meg. Ekkor

$$(x_1, x_2, x_3, x_4, x_5) = (1, 0, 0, 0, 0)$$

állt fenn. Mivel az 1. menetben p meglátogatásakor $B(p) > 0$ teljesült, ezért könnyen ellenőrizhető, hogy p' nem lehetett megmaradó sarokpont az előző iterációban. Ellenkező esetben ugyanis annak bármely 1 értékű $x \in P \vee M$ szomszédjára teljesült volna az $F_3(x) - F_5(x)$ feltételek valamelyike, azaz $B(p) = 0$ adódott volna. Tehát ebben az esetben is ideális pontot töröl az algoritmus.

Végül tegyük fel, hogy p nem kerül törlésre. Akkor $F_1(p') - F_6(p')$ feltételek nem teljesülnek. Ha p a 6(h) ábra maszkjára illeszkedett úgy, hogy

$$(x_1, x_2, x_3, x_4, x_5) = (1, 0, 0, 0, 0),$$

és a következő menetben p már nem törlődik, azaz $B(p) = 0$, akkor p biztosan megmaradó sarokpont volt, hiszen más esetben nem törlődött volna az összes p -vel szomszédos objektumpont. Így ekkor p nem lehetett ideális pont. Ha p nem volt megmaradó sarokpont, viszont eleme volt valamely $\{p, q\}$ kritikus párnak, akkor q törlésre kerül a 2. menet folyamán, különben az m értékével kapcsolatos fenti megállapításaink alapján $m = 1$ lenne az iteráció végére, és emiatt p a következő iteráció 1. menetében törlődne. Azt már beláttuk az előzőekben, hogy ha egy pontot töröl az algoritmus, akkor a törölt pont ideális pont volt, így esetünkben q is ideális pont kellett, hogy legyen. Továbbá mivel q p -vel kritikus párt alkotott, ezért az 5.1. segédétel szerint p nem lehetett ideális pont. \square

5.2. TÉTEL. A javasolt algoritmus bejárásfüggetlen.

Bizonyítás. Az 5.1. tétel szerint az 1. iterációban bejárési sorrendtől függetlenül pontosan az ideális pontok fognak törlődni. Így a következő iterációra jellemző kiindulási kép is egyértelműen meghatározott, és ugyanez elmondható az összes

többi iteráció esetén is, tehát az algoritmus végén kapott kimeneti váz is egyértelmű, függetlenül attól, hogy milyen sorrendben járjuk be a pontokat. \square

5.3. Definíció. Legyen T egy vékonyító algoritmus, p a T kimeneti képmátrixában szereplő tetszőleges objektumpont, $X(p)$ pedig jelöljön egy p -re vonatkozó feltételt. Azt mondjuk, hogy T maximálisan vékonyít az $X(p)$ feltétel szerint, ha bármely bemenet esetén a T által előállított vázban nem szerepel olyan 2×2 -es négyzet alakú objektumrészlet, melyben valamelyik p pixel határpont és teljesül $X(p)$.

5.3. TÉTEL. A javasolt algoritmus maximálisan vékonyít a $(C(p) = 1) \wedge (B(p) \leq 6)$ feltétel szerint.

Bizonyítás. Tegyük fel, hogy marad olyan 2×2 -es négyzet alakú objektumrészlet az algoritmusunk kimeneti képén, amelynek valamely p pontjára teljesül a $(C(p) = 1) \wedge (B(p) \leq 6)$ feltétel. Ezen egyik pont sem lehet ideális pont, különben az eljárás még nem állt volna meg az 5.1. tétel szerint. Ha az objektumrészletünk valamely pontja eleme kritikus párnak és nem γ -pont, akkor az 5.2. segédttétel alapján biztosan lenne még olyan pont, amely ideális, azaz törlésre kerülne, de akkor az előző esethez hasonlóan megint ellentmondásra jutnánk. Ha az objektumrészlet valamely pontja γ -pontként szerepelne egy kritikus párban, akkor pedig a kritikus pár másik eleme kell, hogy α -pont vagy β -pont legyen, különben lenne olyan p pont a kapott képen, amelyre $G_4(p)$ teljesül, vagyis p ideális pont lenne. Viszont ez az eset sem lehetséges az 5.2. segédttétel alapján, mivel akkor is kellene még lennie ideális pontnak a képen. Következésképpen a kimeneti képen nem szerepel kritikus pár. Az említett objektumrészletben valamely p -re $C(p) = 1$, és p nem eleme kritikus párnak. Egy ilyen p csak abban az esetben nem lesz ideális pont, ha megmaradó sarokpont, de ekkor könnyen látható, hogy ebben az objektumnégyesben a p -n kívüli három objektumpont mind ideális pont. Mivel minden lehetséges esetben azt kapjuk, hogy maradt még ideális pont a kimeneti képen, ezért ellentmondásra jutottunk a kezdeti feltevésből, így a tételt bebizonyítottuk. \square

5.4. TÉTEL. A javasolt algoritmus topológia-megőrző.

Bizonyítás. Az iterációk 2. menetében a 2.1. tétel alapján egyértelmű, hogy csak egyszerű pontokat törölünk. Ha az 1. menetben két egymással 4-szomszédos p, q pont kerül törlésre, akkor az 5.1. tétel szerint mindkettő ideális pont, így az 5.2. segédttétel szerint nem alkothattak kritikus párt az előző iteráció elején. Tehát akár p -t, akár q -t töröljük előbb, a 3.1. tétel alapján $C(p) = 1$ vagy $C(q) = 1$ teljesül a másodikként sorra kerülő pontra, ezért ekkor is csak egyszerű pontok kerülhetnek törlésre a 2.1. tételből adódóan.

A 2.2. definíció szerint tehát csak olyan pontokat töröl az algoritmus, amelyek nem változtatják meg a kép topológiáját. \square

6. Eredmények

A javasolt algoritmust számos bináris képre teszteltük és megállapíthattuk, hogy eljárásunk letisztult vázkijelölést eredményez. Módszerünk alkalmazhatóságát most hat tesztképen mutatjuk be (8–13. ábra). Eredményeinket összevetettük a közelmúltban közölt AK^2 algoritmus [2] által kivont vázakkal (8–13. ábra). A hat tesztképre a javasolt és az AK^2 algoritmus futási idejét az 1. táblázat foglalja össze.

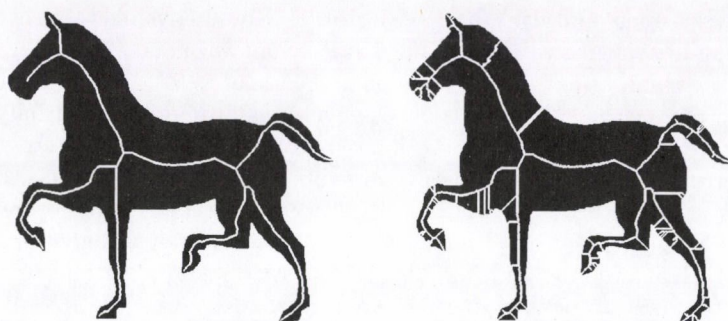
Módszerünk és az AK^2 algoritmus összevetése alapján megállapíthatjuk, hogy az igényelt iterációk számát tekintve nincs közöttük számottevő eltérés, de a javasolt algoritmus egyrészt jelentősen gyorsabb, másrészt pedig nem eredményez felesleges, zavaró vázágot (1. táblázat, 8–13. ábra). A két eljárás összevetésénél megemlítendő még, hogy az általunk javasolt maximálisan vékonyít (vagyis egy pont vastag vázat eredményez), míg az AK^2 algoritmus által produkált vázakban sűrűn előfordulnak két pont vastagságú vonalrészletek.



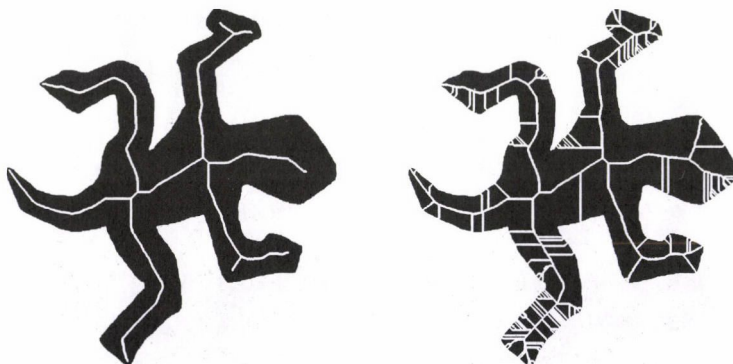
8. ábra. Eredmények egy 2110×880 méretű tesztképre. A javasolt algoritmus (bal) és az AK^2 eljárás (jobb) vázközelítést rávetítettük a kiindulási képre.



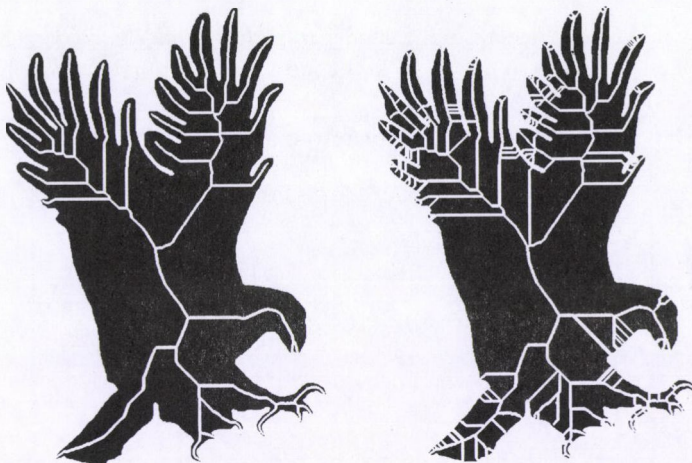
9. ábra. Eredmények egy 1000×1000 méretű tesztképre. A javasolt algoritmus (bal) és az AK^2 eljárás (jobb) vázközelítést rávetítettük a kiindulási képre.



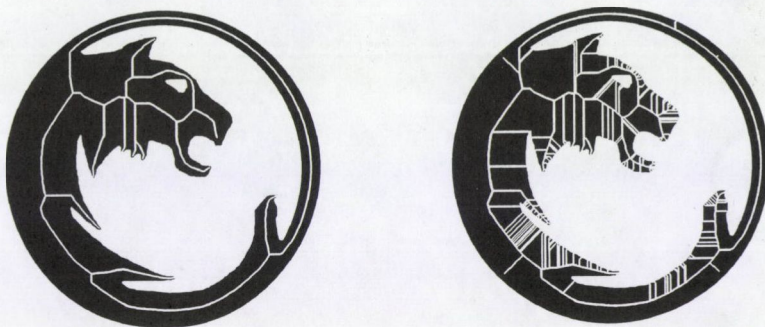
10. ábra. Eredmények egy 524×524 méretű tesztképre. A javasolt algoritmus (bal) és az AK^2 eljárás (jobb) vázközelítést rávetítettük a kiindulási képre.



11. ábra. Eredmények egy 992×1000 méretű tesztképre. A javasolt algoritmus (bal) és az AK^2 eljárás (jobb) vázközelítést rávetítettük a kiindulási képre.









12. ábra. Eredmények egy 644×906 méretű tesztképre. A javasolt algoritmus (bal) és az AK^2 eljárás (jobb) vázközelítést rávetítettük a kiindulási képre.



13. ábra. Eredmények egy 1004×1004 méretű tesztképre. A javasolt algoritmus (bal) és az AK^2 eljárás (jobb) vázközelítést rávetítettük a kiindulási képre.

1. táblázat. A javasolt és az AK^2 algoritmus futási idejének összehasonlítása hat tesztképre. Mindkettő eljárást kétprocesszoros Intel(R) Core(TM)2 6300 (1.86GHz) PC-n futtattuk.

| Tesztkép | Képméret | Iterációszám | | Futási idő (sec) | |
|---|-------------|--------------|----------|------------------|----------|
| | | AK^2 | Javasolt | AK^2 | Javasolt |
|  | 2110 × 880 | 310 | 311 | 42.73 | 20.68 |
|  | 1000 × 1000 | 173 | 174 | 11.82 | 4.25 |
|  | 524 × 524 | 96 | 97 | 1.93 | 0.62 |
|  | 922 × 1000 | 151 | 152 | 9.28 | 3.55 |
|  | 644 × 906 | 179 | 180 | 7.5 | 4.18 |
|  | 1004 × 1004 | 164 | 165 | 11.65 | 5.13 |

Köszönetnyilvánítás

Ezúton szeretnénk köszönetet mondani Gilles Bertrand és Michel Couprie francia kutatóknak (Institut Gaspard-Monge, Laboratoire A2SI, Groupe ESIEE, Noisy-le-Grand), hogy rendelkezésre bocsátották az AK^2 algoritmusuk implementációját.

Hivatkozások

- [1] ARCELLI, C.: *Pattern thinning by contour tracking*, Computer Graphics and Image Processing **17**, (1981) 130–144.
- [2] BERTRAND, G., COUPRIE, M.: *New 2D parallel thinning algorithms based on critical kernels*, in Proc. 11th Int. Workshop Combinatorial Image Analysis, IWCIA 2006, Lecture Notes in Computer Science 4040, Eds. Flach, B., Eckardt U., Knauer, U., Polthier, K., Springer, (2006) 45–59.
- [3] BLUM, H.: *A transformation for extracting new descriptors of shape*, in Models for the Perception of Speech and Visual Form, Ed. Wathen-Dunn, W., MIT Press, Cambridge, (1967) 362–380.
- [4] GONZALES, R.C., WOODS, R.E.: *Digital Image Processing, 3rd Edition* (Prentice Hall, 2008).
- [5] HALL, R.W.: *Parallel connectivity-preserving thinning algorithm*, in Topological Algorithms for Digital Image Processing, Machine Intelligence and Pattern Recognition 19, Eds. Kong, T.Y., Rosenfeld, A.: North-Holland, (1996) 145–179.
- [6] HILDITCH, C.J.: *Linear skeletons from square cupboards*, in Machine Intell. Eds. Meltzer B., D. Michie, D., New York Amer. Elsevier, (1969) 403–420.
- [7] IWANOWSKI, M., SOILLE, P.: *Order independence in binary 2D homotopic thinning*, in Proc. 13th International Conference on Discrete Geometry for Computer Imagery, DGCI 2006, Lecture Notes in Computer Science 4245, Eds. Kuba, A., Nyúl, L., Palágyi, K. Springer, (2006) 592–604.
- [8] KONG, T.Y., ROSENFELD, A.: *Digital topology: Introduction and survey*, Computer Vision, Graphics, and Image Processing **48**, (1989) 357–393.
- [9] KWOK, P.C.K.: *A thinning algorithm by contour generation*, Communications of the ACM **31**, (1988) 1314–1324.
- [10] LAM, L., LEE, S.-W., SUEN, C.Y.: *Thinning methodologies – A comprehensive survey*, IEEE Trans. Pattern Analysis and Machine Intelligence **14**, (1992) 869–885.
- [11] MARCHAND-MAILLET, S., SHARAIHA, J.M.: *Binary Digital Image Processing. A Discrete Approach*, (Academic Press, 2000).
- [12] RANWEZ, V., SOILLE, P.: *Order independent homotopic thinning for binary and grey tone anchored skeletons*, Pattern Recognition Letters **23**, (2002) 687–702.
- [13] ROSENFELD, A.: *A characterization of parallel thinning algorithms*, Information and Control **29**, (1975) 286–291.
- [14] SUEN, C.Y., WANG, P.S.P., (EDS.): *Thinning Methodologies for Pattern Recognition*, Series in Machine Perception and Artificial Intelligence - Vol. 8 (World Scientific, 1994).

(Beérkezett: 2008. augusztus 8.)

KARDOS PÉTER, NÉMETH GÁBOR ÉS PALÁGYI KÁLMÁN

Szegedi Tudományegyetem, Képfeldolgozás és Számítógépes Grafika Tanszék

6720, Szeged, Árpád tér 2.

{pkardos, gnemeth, palagyi}@inf.u-szeged.hu

ORDER-INDEPENDENT SEQUENTIAL THINNING

PÉTER KARDOS, GÁBOR NÉMETH AND KÁLMÁN PALÁGYI

Skeletons are region-based shape descriptors which summarize the general form of objects/shapes. An illustrative definition of the skeleton is given using the prairie-fire analogy: the object boundary is set on fire and the skeleton is formed by the loci where the fire fronts meet and extinguish each other. Thinning is a frequently used method for making an approximation to the skeleton in a topology-preserving way. It is based on a digital simulation of the fire front propagation: the border points of a binary object that satisfy certain topological and geometric constraints are deleted in iteration steps. The entire process is then repeated until only an approximation to the skeleton is left. Sequential thinning algorithms use boundary tracking and delete the actual border point if it satisfies the deletion condition. The skeletons produced by those algorithms generally depend on the applied traversing strategy. A new 2-dimensional sequential thinning algorithm is presented. It is topology-preserving and order independent (i.e., it produces the same skeleton for any traversing strategies).

Keywords: skeleton, thinning, digital topology, topology preservation, order independence.



A FIZIKAI MATEMATIKA LEGÚJABB EREDMÉNYEI MINT A KÖZGAZDASÁG-TUDOMÁNY LEHETSÉGES VIZSGÁLATI ESZKÖZEI

MÓCZÁR JÓZSEF¹

A közgazdaság-tudomány számos problémája a fizika analóg modelljeinek segítségével nyert megoldást. A közgazdászok körében erőteljesen megoszlanak a vélemények, hogy a közgazdasági modellek mennyire redukálhatók a fizika, vagy más természettudományok eredményeire. Vannak, akik pontosan ezzel magyarázzák, hogy a mai mainstream közgazdasági elmélet átalakult alkalmazott matematikává, ami a gazdasági kérdéseket csak a társadalomtudományi vonatkozásaitól eltekintve képes vizsgálni. Mások, e tanulmány szerzője is, viszont úgy vélekednek, hogy a közgazdasági problémák egy része, ahol lehetőség van a mérésre, jól modellezhetők a természettudományok technikai arzenáljával. A másik része, amelyekben nem lehet mérni, s tipikusan ilyenek a társadalomtudományi kérdések, ott sokkal komplexebb technikákra lesz szükség. E tanulmány célkitűzése, hogy felvázolja a fizika legújabb, az irreverzibilis dinamika, a relativitáselmélet és a kvantummechanika sztochasztikus matematikai összefüggéseit, amelyekből a közgazdászok választhatnak egy-egy probléma megfogalmazásában és megoldásában. Például az *időoperátorok* pontos értelmezése jelentős fordulatot hozhat a makroökonómiai elméletekben; vagy az eddigi statikus egyensúlyi referencia pontokat felválthatják a dinamikus, időben változó *sztochasztikus egyensúlyi referencia függvények*, ami forradalmian új megvilágításba helyezhet számos társadalomtudományi, s főleg nemegyensúlyi közgazdasági kérdést. A termodinamika és a biológiai evolúció fogalmait és definícióit Paul A. Samuelson (1947) már adaptálta a közgazdaságtanban, viszont a kvantummechanika legújabb eredményeit, az időoperátorokat stb. nem érintette. E cikk azokat a legújabb fizikai, kémiai és biológiai matematikai összefüggéseket foglalja össze, amelyek hasznosak lehetnek a közgazdasági modellek komplexebb megfogalmazásához.

1. Bevezetés

„Az idő, az elfelejtett dimenzió.” Ilya Prigogine (1980).

¹A szerző köszönetet mond Martinás Katalinnak és Ván Péternek a tanulmányhoz fűzött hasznos észrevételeikért. Külön köszönet illeti meg Matolcsi Tamást és a névtelen lektort a cikk szerkesztéséhez és a tárgyi pontosításokhoz adott javaslataikért. Az esetleges hibákra és tévedésekre a szokásos megjegyzés vonatkozik.

Ezzel a meglepő kijelentéssel kezdi Ilya Prigogine 1980-ban megjelent *From Being to Becoming* című könyvét. Mindezt azzal magyarázza, hogy a dinamika leírásában, egészen napjainkig, legyen az klasszikus vagy kvantum, az idő korlátozott módon lép be, olyannyira, hogy autonóm esetben, az időtől expliciten független állapotegyenleteik, mind a Hamilton-függvény, mind a Schrödinger-operátor invariánsak az idő megfordítására, a $t \rightarrow -t$ -re. Ezek azonban a valóságos esetek elenyésző részének felelnek meg, tehát nem igazán alkalmasak általános elvek megfogalmazására.²

A legkorábban, 1754-ben d'Alembert jegyezte meg az *Enciklopédiájában*, hogy az idő csupán csak geometriai paraméterként jelenik meg a dinamikában. Nála is tovább ment Lagrange (1796), amikor a dinamikát négydimenziós geometriának nevezte. Nézetük szerint a múlt és a jövő ugyanazt a szerepet játsszák. Az atomok és a részecskék által követett világvonalak vagy trajektóriák alkotják az univerzumunkat, amelyek a jövő és a múlt felé egyaránt húzhatók. Vagyis az összes energetikus feltette, hogy minden jelenség tökéletesen *reverzibilis*, és így az egyensúly nem függ az időtől. Jóllehet, közelítéseiket csak autonóm differenciálegyenletekben fogalmazták meg; tehát nincs is igazából tisztázva, hogy mi az egyensúly, és hogyan is függhetne az időtől?

Ez a világszemlélet a nyugati tudomány eredetében gyökerezett, amelyet a 20. század elején csaknem egyhangúlag elfogadott a tudományos közösség. A milétoszi iskola, amelynek az egyik legillusztrisabb képviselője Thálesz volt, bevezetett egy fogalmat az eredeti anyagra, amely szorosan kapcsolódik az anyag konzerválásának koncepciójához. Thálesz szerint egyedül a szubsztancia, mint pl. a víz alkotja az eredeti anyagot. Az összes változásnak a fizikai jelenségekben, mint pl. a növekedés vagy a csökkenés, illúzióknak kell lenniük. A 20. század elején ezt a statikus szemléletet elfogadták az összes tudomány képviselői, a klasszikus (és kvantum-) mechanika képviselői pedig jobbra csak autonóm differenciálegyenletekben gondolkodtak. Ez olyan erősen hatott a tudósokra, hogy még maga Einstein is bevezetett egy kozmológiai állandót³ az általános relativitáselméletbe, hogy továbbra is statikus legyen a világegyetem modellje. (Bár ezt később élete legnagyobb hibájának nevezte.)

E tekintetben a megmaradás elve a döntő, mert definiálja az időbeni azonosságot. Amikor ezt a metaforát átemelték a társadalomtudományba, kiderült, hogy az egyensúlyra itt is igaz, hogy „ami elmúlt, az elmúlt”; így gyakorlatilag figyelmen kívül lehetett hagyni, hogy hogyan működik a piac „valós időben”. Csak a feltételezett végső kifejeletnek szenteltek figyelmet, ami a metaforának éppen a nem statikus voltát mutatja.

Azonban a fizikában és a kémiában is ma már közismert, hogy egy olyan elmélet, amelyben a múlt és a jövő ugyanazt a szerepet játsszák, nem alkalmas minden jelenség leírására. Például, amikor két folyadékot egy kémcsőbe öntünk, általában

²Ráadásul a közgazdasági dinamikában azért sem lehet megfordítani az időt, mert a hasznossági függvény diszkonttényezője monotonitást váltana, és a halálozási kockázat esetén a halott életre kelne.

³A téridőnek (ld. később) immanens tágulási hajlamot tulajdonít.

homogén elegy jön létre. Ebben a kísérletben az idő iránya lényeges. A dinamikus szemlélet a 20. század második harmadától kezdve terjedt el végérvényesen, csaknem valamennyi tudományterületen, ahol az idő már lényeges szerepet játszik. Ezzel az evolúciós fogalom centrális szerepet kapott a fizikai univerzum megértésében, amelyet a fizikában a termodinamika második törvényén, a híres entrópia növekedésének törvényén keresztül vezettek be.

A termodinamika második törvénye megfogalmazása óta hangsúlyozta az irreverzibilis folyamatok egyértelmű szerepét. Valószínűleg a homályos megfogalmazása volt az egyik oka annak, hogy a termodinamika alkalmazása nagyon gyorsan csak az egyensúlyra korlátozódott, a termodinamikai evolúció végső állapotára. Az 1980-as években következett be egy teljes szemléletváltozás, amikor is kezdték megérteni, hogy az irreverzibilis folyamatok konstruktív szerepet töltenek be a fizikai világban, és ezért *a nemegyensúlyi állapotok vizsgálata is jelentős az evolúció tanulmányozásában*. Ennek természetesen előfeltétele volt, hogy a dinamikai rendszer elégségesen „komplex” legyen, hiszen senki sem vár termodinamikai viselkedést például az egyszerű harmonikus oszcillátor esetében. Prigogine megmutatja, hogy a nemegyensúlyi termodinamika⁴ vezet a disszipatív struktúrákhoz, amelyek lényegesek a nemegyensúlyi világban, az életterünkben végbemenő koherencia és szerveződés megértéséhez. Ugyanez a paradigmaváltás figyelhető meg ebben az időszakban a közgazdasági elméletben is: sorra születtek meg a nemegyensúlyi modellek, amelyek stacionárius állapotai az ismert egyensúlyokat eredményezték. *A nemegyensúlyi mikroszkópikus elméletek azonban még nem teljesen kidolgozottak, ezért a közgazdasági elméletek fizikai redukcionizmusában egyféle vákuum keletkezett, ami a felhasználható eredmények elmaradásának is tulajdonítható.*

Klasszikus értelemben a termodinamika második törvénye fejezi ki a molekuláris rendezetlenség növekedését. Amint azt Ludwig Boltzmann⁵ megfogalmazta, és később meg is mutatjuk, *a termodinamikai egyensúly felel meg a maximális „valószínűség” állapotának*. Az evolúció alapértelmezése azonban ennek éppen az ellenkezője a biológiában és a szociológiában, inkább a komplexitás magasabb szintjeire történő transzformációkat írja le. Érzékelhetjük tehát, hogy az idő jelentése különböző a mozgásban, a dinamikában, az irreverzibilitáshoz kapcsolódásában, a termodinamikában, a történelemben, a biológiában és a szociológiában. Prigogine elsőként teszi fel a kérdést, hogy hogyan lehet összekapcsolni az idő ezen különböző fogalmait, hogy egy koherens világszemlélethez jussunk. A választ az idő

⁴ A nemegyensúlyi termodinamikába ad bevezetést Ván (2005) egyetemi jegyzete, miközben a termodinamika különböző egyensúlyi és nemegyensúlyi elméleteit is ismerteti. Ilyenek például a nemegyensúlyi fenomenologikus kontinuum elmélet (irreverzibilis termodinamika), az egyensúlyi statisztikus homogén elmélet (statisztikus fizika), vagy az egyensúlyi fenomenologikus homogén elmélet (egyensúlyi termodinamika vagy termostatika). Az irreverzibilitás közgazdaságtanba történő bevezetésére Martinás (2000) tanulmánya tesz kísérletet.

⁵ Boltzmannt (1844–1906) olyannyira lenyűgözte a termodinamika második törvénye, hogy virtuálisan teljes karrierjét annak megértésének és interpretálásának szentelte. Ebben látta ugyanis Darwin evolúciós elméletének fizikai magyarázatát, s legfőbb ambíciója volt, hogy az anyag evolúciójának „Darwinává” legyen, vagyis kidolgozza az *anyag evolúciójának mechanikai elméletét*.

szerepének újraértékelésében látja. Az irreverzibilis folyamatokat éppoly valósaknak tekinti, mint a reverzibiliseket, amelyek mélyen a dinamikában gyökereznek és a legvilágosabban biológiai szinten jelennek meg. Az irreverzibilitás ott kezdődik, ahol a klasszikus vagy a kvantummechanika alapvető fogalmai (mint pl. a trajektóriák vagy a hullámfüggvények) megszűnnek megfigyelhetőnek lenni. Prigogine itt egy olyan mikroszkópikus formalizációra hívja fel a figyelmet, amely a klasszikus és a kvantum-mechanika szokásos formalizációi mögé terjed, és explicite az irreverzibilis folyamatok szerepét játssza. E formalizáció teszi lehetővé, hogy a fizikai rendszerekre tett megfigyeléseink sok aspektusát a biológiaiakéhoz kapcsoljuk. A szándék itt nem az, hogy a fizikát és a biológiát egyetlen sémára „redukáljuk”, hanem, hogy világosan definiáljuk a leírás különböző szintjeit és megadjuk azokat a feltételeket, amelyek lehetővé teszik a különböző szintek közötti közlekedést.

Míg a klasszikus fizika az euklideszi geometrián alapul, a relativitáselmélet eredményei szorosan kapcsolódnak a geometriai fogalmak kiterjesztéseihez. Az embriológusok a *mezőelméletet* használják a morfogenezis komplex jelenségeinek ábrázolására. Egy fejlődő embrió modellezése rendkívül nehéz: a sejtek száma változik, új szövetek és szervek keletkeznek, amelyek konfigurációja fontos alkotórésztévé válik az állapotnak. A biológiai térben minden egyes esemény egy pillanat alatt következik be, és egy olyan régióban, ami lehetővé teszi, hogy a folyamatot teljes egészében koordináták közé helyezzük. Ez a tér funkcionális és nem geometriai. Az ökológiai állapottér gyakran a populáció olyan listája, amely a fajok adott halmazát tartalmazza. A standard geometriai tér, az euklideszi tér invariáns az eltolásokra és a forgatásokra. A biológiai és az ökológiai térben viszont az események időben és térben elhelyezett folyamatok és nem csupán trajektóriák. Egészen közel kerülünk a kozmosz arisztotelészi nézetéhez, amely szembeállította az isteni világot és az örökkévaló trajektóriákat a 'szublunaris' természetű (vagyis a valóságos) világgal, amelynek leírásában nyilvánvalóan meghatározó szerepet játszottak a biológiai megfigyelések.

Eudoxosz, Ptolemaiosz, Kopernikusz és Brahe geometriai modelljei több-kevesebb sikert jelentettek a Naprendszer dinamikájának leírásában, de ezekből hiányzott egy olyan vezérelv, ami biztosította volna a modellek szisztematikus felépítését. Kopernikusz heliocentrikus modellje óriási filozófiai jelentősége mellett sem garantált lényegesen pontosabb előrejelzéseket, mint Ptolemaiosz geocentrikus modellje.

Bár Arisztotelész biológiai nézeteinek alkalmazásai a fizikára katasztrofális következményekkel jártak, a modern bifurkációs és instabilitási elméletek azt mutatják, hogy a két fogalom, a geometriai és a szerves funkcionális világ nem inkompatibilisek. A klasszikus, gyakran Galileinek tulajdonított nézet a világot egy objektumnak tekintette, s a fizikai világot külső szemlélőként írta le. Az élő organizmusok az egyensúlytól messze lévő objektumok, amelyeket az instabilitások szeparáltak az egyensúly világától, és amelyek szükségszerűen nagyok, makroszkópikus objektumok. Az anyag egy koherens állapotát követelik meg azért, hogy előállítsák a komplex biomolekulákat, amelyek lehetővé teszik az élet örökkévalóságát.

Prigogine próbálkozása az ún. „második idő” fogalmának bevezetésére, csak egy a sok közül. Ezek a kísérletek abból indulnak ki, hogy a mozgással kapcsolatos idő csak az első aspektus, amely konzisztensen belefoglalható a klasszikus fizika és a kvantummechanika kereteibe. A *második idő* megfogalmazása, amint később bemutatjuk a legelső kísérleteket, a fluktuációkban gyökerezik, mélyen a mikroszkopikus dinamika szintjén. Ebben a közelítésben az idő többé nem egy egyszerű paraméter, mint a klasszikus, vagy a kvantummechanikában, hanem egy *operátor*, azokhoz hasonló, amelyek leírják a mennyiségeket a kvantummechanikában. Az operátorok azért szükségesek, hogy leírassuk a mikroszkopikus szint váratlan komplexitásait. A relativitáselméletben az idő nem különül el a tértől, és nem is független tőle, hanem vele együtt a *tér-időnek* nevezett rendszert alkotja.

A klasszikus fizika, még a kvantummechanikával és a relativitáselmélettel együtt is, a különböző rendszerek időbeli evolúcióinak viszonylag gyenge modelljeit adta. A fizika determinisztikus törvényei, amelyek egy ponton az egyedüli elfogadható törvények voltak, ma nagyon nagy egyszerűsítéseknek tűnnek, valósággal az evolúció karikatúrái. Mind a klasszikus, mind a kvantummechanikában, ha egy rendszer állapota elfogadható pontossággal ismert volt egy adott időpontban, akkor a jövő (és a múlt is) előrejelezhető volt, legalábbis elvileg. Az elméleti keret predeterminálta, hogy a jelen, bizonyos értelemben, már tartalmazza a múltat és a jövőt. Prigogine kutatásai mindezt cáfolják. Még a fizikában is, épp úgy, mint a biológiában, csak különböző lehetséges scenáriók jelezhetők előre.

Az integrálható rendszerek egyensúlyának megkeresésében jelentkező nehézségek miatt James Clerk Maxwell és Boltzmann a dinamikai rendszerek egy másik típusát kezdték el vizsgálni, nevezetesen azokat, amelyeket ma *ergodikus rendszereknek* nevezünk. A termodinamikai egyensúly problémájának közvetlen bizonyítására bevezették az ergodikus hipotézist. Ez a következőket állítja: ha a rendszer marad a tényleges mozgásállapotában, akkor előbb vagy utóbb keresztülmegy minden olyan fázison, amely kompatibilis az energia egyenlettel. Ugyanakkor a matematikusok kimutatták, hogy a trajektóriák nem töltik ki teljesen a „felületet”, de a dinamikai rendszer bármelyik trajektóriája tetszőlegesen közel kerülhet bármely állapothoz (az energiafelület mindegyik pontjához), összhangban a *kvázi-ergodikus hipotézissel*.

Ennek illusztrálására tekintsünk egy egyszerű példát, mégpedig a kétdimenziós egységnyegyzeten végbemenő mozgást, amit az alábbi egyenletekkel adunk meg:

$$\frac{dp}{dt} = \alpha \quad \text{és} \quad \frac{dq}{dt} = 1.$$

A rendszer megoldásai az induló feltételekkel együtt:

$$\begin{aligned} p(t) &= p_0 + \alpha t, \\ q(t) &= q_0 + t. \end{aligned}$$

Az ezekből nyerhető fázistrajektória:

$$p = p_0 + \alpha (q - q_0).$$

A trajektória főbb karakterisztikumai α értékeitől függnnek, amelyekre két eset különböztethető meg. Ha α racionális szám, mondjuk $\alpha = m/n$, ahol m és n relatív prímek (egészek), akkor a trajektória periodikus lesz, és $T = n$ periódus után ismételni fogja önmagát. A rendszer ebben az esetben nem ergodikus. Ha viszont α irracionális, akkor a trajektória kielégíti a kvázi-ergodikus hipotézis feltevését. Az egységnégyzet minden egyes pontjához tetszőlegesen közel fog kerülni, de nem tölti ki teljesen a négyzet felületét.

Az a gondolat, hogy a dinamikai rendszerek csak általában ergodikusak, először Kolmogorov (1954) tanulmányában jelent meg. Kimutatta, hogy interaktív dinamikai rendszerek nagy osztályaira lehet szerkeszteni olyan periodikus pályákat, amelyek az ergodikus felület egy alterére (invariáns tóruszokra) korlátozóztak. Más vizsgálatok is ismertek, amelyek gyengítették az ergodikus rendszerek univerzalitásába vetett hitet. Például Enrico Fermi, John Pasta és Stanislaw Ulam nemharmonikus oszcillátorok összekapcsolt láncainak viselkedését vizsgálták. Azt várták, hogy a rendszerük gyorsan elér egy termális egyensúlyt. Ehelyett periodikus oszcillációkat kaptak, mégpedig különböző normál móduszú energiában. Kolmogorov elméletét kiterjesztette Arnold és Moser, és ez vezetett az ún. *KAM-elmélet*hez. Ennek az új elméletnek talán az a legérdekesebb aspektusa, hogy *függetlenül az ergodicitástól, a dinamikai rendszerek véletlenszerű mozgásra vezethetnek*⁶.

A gravitációt a tudomány mai állása szerint nem vonzó erőként, hanem térgörbületként fogjuk fel. A tér és idő külön nem létezik, egységes egészet alkotnak. Ezen belül minden pontszerű testnek értelmezhető a sajátideje. Egy megfigyelő („rendszer”) nem más, mint pontszerű testek összessége, minden pontnak létezik sajátideje, de ezek a sajátidők általában egymástól függetlenek, nem állnak össze egységes rendszeridővé. Rendszeridőt csak szinkronizációval tudunk létrehozni, de a szinkronizáció és a sajátidők általában egymástól függetlenek.

Jelen világunkat különböző szinteken vizsgálva, más és más törvényszerűségeit ismerhetjük meg. A mikrovilág szintjén a kvantumfizika törvényei a mérvadók, a makrovilág szintjén pedig a kémia és a biológia törvényei, míg az Univerzum szintjén az Einstein-féle relativitáselmélet az elfogadott. A mai (nyugati) fizikai (közgazdasági) szemlélet alapvető hibája, hogy azt hiszi, hogy a mikroból megérthető a makro. Azt tanítják, hogy a fenti szintek csak a leírásukat tekintve különbözöek, valójában az egész Univerzumunk a mikrovilág elemeiből épül fel. Minél többet tudunk a mikroszkopikus világról, annál többet tudunk az Univerzumból. Hasonló szemlélet uralkodik a közgazdaságtanban is: a makrofolyamatok alakulását a mikroökonómiából kiindulva magyarázzák. Ami a két terület formalizmusát illeti, érdemes itt idéznünk Kornai (1971, 226 o.) könyvéből: „A klasszikus mechanika matematikai apparátusa a makrofizikára alkalmazható, viszont a mikrofizika kvantumos szemléletű. Analóg a helyzet a gazdasági valóságban: a mikroökonómia nagy, aggregált folyamatai jól leírhatók folytonos változókkal, míg a mikroökonómia számos jelensége kvantumos.” A makro ún. mikro értelmezései erősen vitathatók,

⁶A dinamikai rendszerek egy átfogó modern matematikai vizsgálata található Hirsch (1989) kitűnő tanulmányában.

mivel a részecskék együttes viselkedése nem következethet önnön tulajdonságai-ból. Ugyanúgy, mint az egyes emberek fogyasztói magatartási tulajdonságaiból sem aggregálhatók hasonló társadalmi viselkedések, ami elméleti bizonyítást is nyert a híres DSM-tételekben⁷.

2. Idő a fizikában

Az idő az ősrobbanás, a világegyetem keletkezésének pillanatában kezdődött, s bármennyire is furcsán hangzik, az ennél korábbi időpontok teljesen határozatlanok. S minthogy *Albert Einstein* általános relativitáselmélete értelmében a világegyetemenk is vége lesz egyszer, ezért az idő, a negyedik dimenzió véges kiterjedésű. *Stephen W. Hawking* és *Robert Penrose* azonban bebizonyították, hogy az általános relativitáselmélet a kvantummechanika határozatlansági tételével (ld. később) együtt biztosítja, hogy az idő a végessége mellett mégis határtalan legyen, csak úgy, mint a tér⁸. Az ún. *szingularitási tételükkel* bebizonyították azt is, hogy az idő kezdete egyetlen, végtelen sűrűségű és a téridőben végtelen görbületű pont volt. Mielőtt rátérnénk ennek mélyebb következményeire, érdemes röviden kronológiailag is áttekinteni az idő fogalmának különböző értelmezéseit a természettudományokban.

Arisztotelész számára a fizika a folyamatok, a természetben különböző időpontokban megfigyelhető változások tudománya volt. A modern fizika megalapozói, köztük Galileo Galilei számára is, az egyetlen változás, ami pontos matematikai képletekben is kifejezhető, a gyorsulás, a mozgásban bekövetkező változás volt. Az időt egy tisztán fizikai jelenség absztrakt paraméterének tekintette, ami lehetővé tette a mozgás kvantifikációját. Newtonnál mindez, mint „abszolút” idő jelenik meg. Azt tanította, hogy két esemény közötti időtartam egyértelműen mérhető, teljesen függetlenül a tértől. Ez vezetett el végül is a klasszikus mechanika fundamentális egyenletéhez, amely a gyorsulást a tömeggel arányosan az erőhöz kapcsolja. A fizikai időt a mozgási egyenletekben megjelenő idővel azonosították.

⁷Elsőként Hugo F. Sonnenschein fejtette ki aggályait két cikkében is (*Sonnenschein* (1972, 1973)), az előbbi követte *Debreu* (1972), majd pedig Rolf R. Mantel cikke (*Mantel* (1974)). Mindegyikük abból indult ki, hogy a piaci keresleti és túlkeresleti függvényeket a fogyasztók hasznosságmaximalizáló tevékenységeinek összegzésével definiálják. A három szerző azt állítja, hogy a piaci keresleti és túlkeresleti függvények, amelyeken a piacsintű mikroökonómia és makroszintű makroökonómia összes „intuitív” állítása nyugszik, nem rendelkeznek azokkal a tulajdonságokkal, mint amilyenekkel a fogyasztói keresleti és túlkeresleti függvények. Egyszerűbben fogalmazva: például, mégha mindenkinek szabályos alakú egyéni keresleti függvénye is van, nem mondhatjuk azt, hogy a piaci keresleti függvény is biztosan szabályos alakú lesz. Csak nagyon speciális esetben várható, hogy a gazdaság úgy viselkedik, mint egy idealizált fogyasztó. Ez valójában romba döntötte a közgazdasági elmélet „mikromegalapozási” megközelítését - azt, hogy az aggregált keresletet és kínálatot a hasznosságmaximalizáló piaci szereplők viselkedéseként írják le. Így lényegében hiábavalóknak bizonyultak azok az elmúlt századi erőfeszítések, hogy az aggregált keresletet a hasznosságmaximalás eredményeként szerepeltessék. (Az Olvasó egy kitűnő áttekintést kaphat ezekről a munkákról *Shafer és Sonnenschein* (1982) tanulmányából.)

⁸Valójában az idő és az energia határozatlansága nem következik a kvantummechanika ma ismert megfogalmazásából (noha a kísérleti eredmények mégis utalnak ilyesmire).

A Lagrange-függvény, illetve a Hamilton-egyenletekből nyert dinamika valójában nem tesz különbséget a jövő és a múlt között. A differenciálegyenletekből nyert trajektóriák gyűjteményével értelmezték a fizikai világot, a természet minden olyan elemi megnyilvánulását, amelyben szerepet játszik az idő.

Itt érdemes megjegyezni, hogy a differenciálegyenletek ilyen irányú vizsgálata termékenyítőleg hatott az absztrakt matematika minden területén: Fourier a hővezetés tanulmányozásakor alkotta meg sorait, Cantort pedig éppen a Fourier-sorok konvergencia problémái vezették el a topológiához és a halmazelmélethez, Lie felfedezte a róla elnevezett csoportokat, és Poincaré kifejlesztette az algebrai topológiát.

Az elméleti fizika ideálja egy olyan stabil világ volt egészen mostanáig, amely elkerüli a valamivé válás folyamatát. Isaac Newton dinamikája, amelyet olyan nagy követői, mint például Pierre Laplace, Joseph Lagrange és Sir William Hamilton tettek teljessé, egy zárt univerzális rendszert fogalmaztak meg, amellyel minden kérdés megválaszolását lehetségesnek tartották. Ha egy kérdésre mégsem tudtak válaszolni, azt definíciószerűen pszeudo problémának tekintették, csak illúzióként kezelték, amelynek nincs semmiféle fundamentális jelentősége. Hosszú ideig úgy tűnt, hogy a dinamika lehetővé teszi a realitások teljes megismerését.

Igy lett a fizika fő célkitűzése a mikroszkopikus szint olyan azonosítása, ami lehetővé teszi a dinamika alkalmazását, s amely alapot ad az összes lehetséges megfigyelhető jelenség magyarázatához. Itt találkozott a klasszikus fizika a görög atomisták programjával, amint azt Démokritosz állította: „Csak az atomok és az űr!”

Ma már tudjuk, hogy a newtoni dinamika a fizikai tapasztalásunknak csak egy részét írja le, az objektumok tömegét pedig grammokban, sebességét pedig a fényénél kisebb sebességben méri. Az is közismert, hogy a klasszikus dinamika érvényességét korlátozzák az univerzum konstansai, mint pl. a h Planck konstans, amelynek értéke az SI rendszerben $6.6260755 \times 10^{-34} Js$ nagyságrendű és a c , a fény sebessége $\sim 3 \cdot 10^{10}$ cm/sec. Az olyan új jelenségek, mint pl. a kis objektumok, vagy a hypersűrűségű neutron csillagok, vagy a fekete lyukak tanulmányozásához a newtoni dinamikát felváltották a kvantummechanikával (amely véges értékű h -t vesz figyelembe) és a relativitáselmélettel (amely már tartalmazza a c -t is). A relativitáselméletben nagy sebességek esetén a tér és az idő összefonódik, míg a kvantummechanikában a megfigyelés befolyásolja a tárgy állapotát. Azonban a dinamikának ezen új formái is örökölték a newtoni fizika ideálját, a statikus univerzumot, a létezés univerzumát a valamivé válás nélkül. Az abszolút idő azonban megszűnik, minden megfigyelő számára más mérték szerint múlik az idő. A csillag gravitációs tere hatására a felszínen tartózkodó személy számára másként telik az idő, mint az űrbeli megfigyelő számára.

A fenti értelmezésekben azt láttuk, hogy a dinamika olyan folyamatokat ír le, amelyekben az idő iránya irreleváns. Vannak azonban olyan helyzetek is, amikor az idő iránya valóban lényeges szerepet játszik. A legújabb kutatások szerint (ld. például Hawking (1998)) az időnek legalább háromféle irányítottága van. Az első, ami tanulmányaink szempontjából most a legfontosabb, a termodinamikai irány, ami mentén a rendezetlenség vagy az entrópia nő. A második a pszichológiai irány,

ami arrafelé mutat, amerre érzékeink szerint halad az idő. Ebben az irányban a múltra emlékeztünk és nem a jövőre. Hawking (1998) megmutatta, hogy e két időirány egyfelé mutat. A harmadik a kozmológiai irány, amiben a világegyetem tágul és nem zsugorodik. Vizsgáljuk meg most alaposabban az első irányt!

Ha a makroszkópikus tér egy részét melegítjük és ezután termálisan izoláljuk a tér-részt, megfigyelhetjük, hogy a hőmérséklet bizonyos idő eltelte után egyforma lesz az egész térben (*Laplace-démon*). E folyamatban az idő nyilvánvaló „egyirányúságot” mutat. A termodinamika második törvénye összegzi e folyamat jellemző tulajdonságait. Rudolf Clausius olyan izolált rendszereket vizsgált, amelyek sem anyagot, sem energiát nem cserélnek a külvilággal. Ekkor a második törvény implicálja az S függvényt, az entrópia létezését, amely mindaddig monoton nő, amíg el nem éri a maximális értékét a termodinamikai egyensúly állapotában:

$$dS/dt \geq 0.$$

Az entrópia változását, a dS -t a $d_e S$, az entrópiának a rendszer határain keresztüli transzfere, és a $d_i S$, a rendszeren belül termelt entrópia határozza meg, azaz

$$dS = d_e S + d_i S, \quad \text{ahol} \quad d_i S \geq 0.$$

E megfogalmazásban a reverzibilis és az irreverzibilis folyamatok közötti distinkció lényeges. Ugyanis csak az irreverzibilis folyamatok járulnak hozzá az entrópia termeléséhez. Ilyen irreverzibilis folyamatok, például a kémiai reakciók, a hővezetés és a diffúzió. A reverzibilis folyamatok a korlátba ütköző hullámterjedés, amelyen a hullám elnyelése kizárt. A termodinamika második törvénye ekkor azt állítja, hogy az irreverzibilis folyamatok az idő egyirányúságához vezetnek. Az entrópia, másképpen a rendezetlenség növekedése különbözteti meg a jövőtől a múltat, vagyis irányt ad az időnek.

A pozitív időirány tehát az entrópia növekedésével társul. Ez feltételezi egy egészen specifikus tulajdonságokkal rendelkező függvény létezését, amely izolált rendszerben az idő monoton növekvő függvénye. Ahhoz, hogy ezt részletesebben is tárgyalhassuk, tekintsünk egy olyan rendszert, amelynek evolúcióját bizonyos X_i változók írják le, amelyek, például a kémiai elemek koncentrációját mutatják. A rendszer evolúcióját a

$$dX_i/dt = F_i(\{X_i\})$$

alakú rátaegyenletek írják le, amelyekben F_i az X_i komponens átlagos változási sebessége. Természetesen mindegyik komponensre van ilyen egyenletünk. Tegyük fel, hogy az $X_i = 0$ -ra valamennyi reakciósebesség zérus lesz. Ekkor ez a rendszer egyensúlyi pontja. Kérdés, hogy amennyiben a koncentrációk nem zérus értékeiből indulunk ki, a rendszer eljut-e az $X_i = 0$ egyensúlyi állapotba? Vajon tekinthető-e az $X_i = 0$ állapot attraktornak? A kérdés megválaszolásában a Ljapunov-függvény segít. Vegyük a koncentrációk egy bizonyos

$$V = V(X_1, X_2, \dots, X_n)$$

függvényét, és tegyük fel, hogy értéke pozitív a kérdéses értelmezési tartományban és $X_i = 0$ -ban zérus. Ekkor megvizsgálhatjuk, hogy hogyan változik a $V(X_1, X_2, \dots, X_n)$, amint az X_i koncentrációk fejlődnek. Ehhez a függvény idő szerinti deriváltját kell vennünk, vagyis

$$\frac{dV}{dt} = \sum_i \frac{\partial V}{\partial X_i} \frac{dX_i}{dt}.$$

Ljapunov (1908) tétele azt állítja, hogy az egyensúlyi állapot attraktor, ha a dV/dt előjele V előjelének az ellentettje; azaz a mi esetünkben most negatív. Ennek a feltevésnek a geometriai jelentése evidens. Izolált rendszerekre a *termodinamika második törvénye* azt állítja, hogy a Ljapunov-függvény létezik, és hogy az ilyen rendszerekre a *termodinamikai egyensúly a nemegyensúlyi állapotok attraktora*. Ezért hangsúlyozta Max Planck (1930), hogy a termodinamika második törvénye megkülönbözteti a természeti állapotok egyes típusait a többiekétől, ha azok attraktorként viselkednek a többiek számára. Az *irreverzibilitás ennek a vonzásnak a kifejeződése*.

Az irreverzibilis folyamatok molekuláris leírásában az alapvető kérdés az, hogy mit jelent az entrópia növekedése molekulákban kifejezve? A válaszhoz meg kell adnunk az entrópia mikroszkópikus jelentését. Boltzman jegyezte meg elsőként, hogy az *entrópia a molekuláris rendezetlenség mértéke*, és arra a következtetésre jutott, hogy az entrópia-növekedés törvénye egyszerűen a növekvő rendezetlenség törvénye. Tekintsünk például egy konténert, amelyet két egyforma rekeszre osztunk. Azon esetek P száma, amelyek N számú molekulának a két rekeszbe történő rendezését mutatja, egy egyszerű kombinatorikus formulával megadható:

$$P = \binom{N}{N_1} = \frac{N!}{N_1!(N - N_1)!} = \frac{N!}{N_1!N_2!}.$$

Egy bizonyos idő elteltével elérjük az egyensúlyi állapotot, amelyben közel egyenlő számú molekula van mindkét rekeszben, azaz $N_1 \simeq N_2 = N/2$. Könnyen belátható, hogy ebben az állapotban veszi fel P a maximális értékét, és ebben az evolúciós folyamatban P nő. Boltzman az entrópiát a P segítségével következőképpen fejezte ki:

$$S = k \log P,$$

amelyben k Boltzman univerzális konstansa ($1,3806568 \times 10^{-23} \text{ JK}^{-1}$). Az entrópia növekedése a növekvő molekuláris rendezetlenséget fejezi ki, vagyis P növekszik. Ebben az evolúciós folyamatban az induló állapot elfelejthető, a mindenkor aszimmetria, vagyis, hogy az egyik rekeszben több a molekulák száma, idővel mindig megszűnik. Ha P egy állapot „valószínűségét” méri, akkor az entrópia növekedése megfelel a „legvalószínűbb” állapot felé történő evolúciónak. A *valószínűség fogalma az irreverzibilitás molekuláris értelmezésén keresztül lépett be az elméleti fizikába, ami döntő lépés volt a modern fizika történetében*.

A valószínűségi érvelésben még tovább kell lépnünk, ha olyan kvantitatív megfogalmazásokat akarunk nyerni, amelyek leírják, hogy az irreverzibilis folyamatok

hogyan fejlődnek idővel. Tekintsük például a jól ismert bolyongási problémát, a *Brown-mozgás* idealizált, de mégis sikeres modelljét. A legegyszerűbb példában, az egydimenziós bolyongásban, a molekula egy lépést tesz valamelyik irányban a számegegyes szabályos intervallumaiba. Ha a molekula kezdetben az origóban volt, kérdezhetjük, hogy mi a valószínűsége annak, hogy N lépés után az m pontban lesz. Ha annak valószínűségére, hogy a molekula előre, vagy hátra halad-e, feltesszük, hogy $1/2$, akkor

$$W(m, N) = \left(\frac{1}{2}\right) \frac{N!}{[(1/2)(N+m)]! [(1/2)(N-m)]!}. \quad (1)$$

Így a molekulának, hogy az m pontba érkezzon N lépés után, $(1/2)(N+m)$ lépést kell tennie jobbra és $(1/2)(N-m)$ -et balra. Az (1) egyenlet az ilyen különböző sorozatok számát adja megszorozva az N lépés egy tetszőleges sorozatának teljes valószínűségével. (A részletekért lásd Chandrasekhar (1943).)

Kifejtve a faktoriálisokat, egy Gauss-eloszlásnak megfelelő aszimptotikus formulát kapunk:

$$W(m, N) = \left(\frac{2}{\pi N}\right)^{1/2} e^{-m^2/2N}.$$

Felhasználva a $D = (1/2)nl^2$ jelölést, amelyben az l a két hely közötti távolság és n az egységnyi idő alatti elmozdulások száma, a fenti egyenletet a következőképpen is írhatjuk:

$$W(x, t) = \left(\frac{1}{2(\pi Dt)^{1/2}}\right) e^{-x^2/4Dt}, \quad (2)$$

amelyben $x = ml$.

A (2) egyenlettel definiált Brown-mozgást az egységnyi idő alatt jól definiált átmeneti valószínűségekkel is megfogalmazhatjuk. Tekintsük ismét annak a valószínűségét, hogy a Brown-részecske a t időpontban a k helyen van, és jelöljük ezt $W(k, t)$ -vel. Vezessük be az ω_{lk} átmeneti valószínűséget, amely egységnyi idő alatt a két állapot, a k és l közötti átmenet valószínűségét jelöli. Ekkor kifejezhetjük a $W(k, t)$ idő szerinti változását olyan versengésben kifejezve, amely az $l \rightarrow k$ -hoz kapcsolódó nyertesek és a $k \rightarrow l$ -hez kapcsolódó vesztesek között folyik. Ekkor felírhatjuk az alapegyenletet:

$$\frac{dW(k, t)}{dt} = \sum_{l \neq k} [\omega_{lk} W(l, t) - \omega_{kl} W(k, t)].$$

A Brown-mozgásban k a rácson levő helyzetnek felel meg, és az ω_{kl} csak akkor különbözik zérustól, ha k egy egységgel különbözik az l -től. De a fenti egyenlet ennél is sokkal általánosabb. Valójában alapegyenlete a *Markov-folyamatoknak*, amely vezető szerepet játszik a modern valószínűségelméletben. (Lásd Barucha-Reid (1960).)

A *Markov-folyamat* egyik lényeges tulajdonsága, hogy az ω_{lk} átmeneti valószínűségek csak a k és az l állapotokat foglalják magukba. Az átmeneti valószínűség $k \rightarrow l$ -be nem függ attól, hogy mely állapotok voltak „bejárva” a k állapot bekövetkezése előtt. *Ebben az értelemben a rendszernek nincs memóriája.* (Ezért kérdéses, hogy mennyire alkalmas előrejelzések felhasználására.) A Markov-folyamatokat sok fizikai állapot leírására használták fel és felhasználhatók a kémiai reakciók modellezésére is.

Amíg a klasszikus és a kvantumdinamikában a fizika alapvető törvényei időben szimmetrikusak, addig a termodinamikai irreverzibilitás nem az, lényegében a dinamikához hozzáadott egyféle közelítésnek felel meg. Egy gyakran idézett példa Josiah Willard Gibbs (1902) nevéhez fűződik: ha egy csepp fekete tintát pottyantunk egy pohár vízbe és megkeverjük, akkor az elegy szürke színű lesz. Ez a folyamat irreverzibilisnek látszik. De ha követni tudnánk minden egyes molekulát, felismernénk, hogy a mikroszkópikus birodalomban a rendszer heterogén maradt. Az irreverzibilitás csak egy illúzió lenne, amelyet a megfigyelő érzékszerveinek tökéletlensége váltott ki. Igaz, hogy a rendszer heterogén maradt, de a kezdeti makroszkópikus heterogenitási skála mikroszkópikussá lett. Az a nézet, hogy ebben az értelemben az irreverzibilitás csak egy illúzió, nagyon hatásosnak bizonyult, és sok tudós megpróbálta hozzákötni ezt olyan matematikai eljárásokhoz, amelyek irreverzibilis folyamatokhoz vezetnek. Mások, hasonló céllal, a makroszkópikus megfigyelés feltételeit próbálták meg kidolgozni. Ezen kísérletek egyike sem vezetett azonban meggyőző eredményre.

Ugyanakkor nehéz elhinni, hogy bizonyos megfigyelhető irreverzibilis folyamatok, olyanok, mint a viszkózitás, az instabil részecskék számának csökkenése stb. egyszerűen csak illúziók lennének, amelyeket a tudás hiánya, vagy a tökéletlen megfigyelés váltott ki. Mivel az induló állapotot, még a legegyszerűbb dinamikai mozgásban is, csak közelítőleg tudjuk megadni, a mozgás jövőbeli állapotait az idő előre haladtával egyre nehezebb lesz előrejelezni. Az ilyen rendszerekre nem alkalmazhatjuk a termodinamika második törvényét. Alkalmazhatjuk a sok interaktív részecske által formált gázra, de nem alkalmazhatjuk olyan egyszerű dinamikai rendszerekre, mint például a bolygórendszer. Mindebből Prigogine arra a következtetésre jutott, hogy az *irreverzibilitásnak valamilyen kapcsolatban kell lennie a kérdéses rendszer dinamikus természetével.*

Vizsgálták természetesen annak lehetőségét is, hogy talán a dinamika tökéletlen, talán azt kellene kiterjesztteni az irreverzibilis folyamatok befoglalásával. Azonban ez az út sem volt járható, mivel a dinamikus rendszerek egyszerű típusaira mind a klasszikus, mind a kvantum mechanikában az előrejelzések jól beigazolódtak. Elegendő csak megemlíteni az űrkutatás sikereit, ami a dinamikus trajektóriák nagyon pontos számításait követelte meg. Ennek ellenére mostanában ismételten felteszik a kérdést, hogy a kvantum mechanika teljes-e az ún. mérési problémával összefüggésben? Nem kellene-e a mérés irreverzibilitásának befoglalásához új tagot hozzáadni a kvantumrendszerek dinamikáját leíró *Schrödinger-egyenlethez*? Prigogine kutatásai pontosan ezeket a „statikus” dinamika és az irreverzibilitást hangsúlyozó termodinamika közötti kapcsolódásokat tárják fel.

A kvantumelmélet események összegzésén alapuló megfogalmazásában a részecskének nincs egyetlen múltja, minden megengedett pályát befut a téridőben, amelynek mérete véges, de hipotetikusán nincs határa vagy pereme. Minden ilyen pályát *Hawking* két számmal jellemez: az egyik a nagyságára jellemző, a másik a fázisára (a cikluson belüli helyzetére). Ha arra vagyunk kíváncsiak, hogy mekkora valószínűséggel halad át a részecske egy bizonyos ponton, akkor összegeznünk kell mindazon hullámokat, amelyek a pontot érintő pályákhoz tartoznak. Ilyenkor a képzetes idő használatához kell folyamodnunk, vagyis az időt nem valós, hanem képzetes számokkal kell mérnünk. Ennek érdekes hatása van a téridőre: megszűnik a tér és az idő közötti különbség. Az olyan téridőt, ahol az események időkoordinátájának képzetes értéke van, *euklideszi téridőnek* nevezzük. Ebben nincs különbség az idő iránya és a térirányok között. A *valós téridőben* viszont, ahol az eseményeket a valós időkoordináta jellemez, szembeeszkő a különbség: az idő iránya minden pontban a fénykúpon belül van, a térbeli irányok pedig azon kívül.

3. A klasszikus mechanika dinamikája

A klasszikus mechanika a mai elméleti fizika legrégebbi múltra visszatekintő ága, a 20. század tudományos forradalmainak kiinduló pontja, ami végül is a relativitáselmélethez és a kvantummechanikához vezetett. A klasszikus mechanika jórészt abból áll, hogy Newton törvényeiből, például az energiamegmaradásából, olyan differenciálegyenleteket vezetünk le, amelyek sok, véges számú szabadságfokkal rendelkező fizikai rendszerek „mozgástörvényeit” írják le. Franciául a „*rationális*” mechanika kifejezést használják gyakrabban, ezzel is hangsúlyozva, hogy a klasszikus mechanika törvényei éppen a józan ész törvényei. A klasszikus dinamikához rendelt karakterisztikumok közt találjuk a szigorú determinizmus törvényét. A dinamikában alapvető különbséget tesznek a kezdeti feltételek, amelyek tetszőlegesen adhatók meg, és a mozgási egyenletek közt, amelyekből kiszámítható a dinamikus rendszer későbbi (vagy korábbi) állapota. A modern dinamika Johannes Keplernek a bolygók mozgási törvényeivel és Newton „két-test” problémájával született meg. Azonban a dinamika rendkívül bonyolulttá válik mihelyst figyelembe veszünk egy harmadik testet is, például még egy másik bolygót is. Ha a rendszer elégségesen komplex (mint a „három test” problémában), még a rendszer kezdeti állapotának (bármilyen véges pontosságú) ismerete sem teszi lehetővé általában a rendszer viselkedésének előrejelzését hosszabb távra.⁹ Ez a bizonytalanság fennáll még akkor is, amikor a pontosság a kezdeti állapot meghatározásában tetszőlegesen nagyra válik. Lehetetlen, még csak elvben is, például azt tudni, hogy

⁹Itt külön kell választani az elvet és a gyakorlatot. A mechanikai mozgásokat (pl. az égitestekét) olyan differenciálegyenlet írja le, amelynek adott kezdeti feltételek mellett egyértelmű megoldása létezik. Csakhogy elég bonyolult differenciálegyenlet esetén (pl. a háromtest problémáiban) a megoldást nem tudjuk analitikus formában előállítani.

a Naprendszer, amelyben élünk, stabil-e örök időkre?¹⁰ Ezek a megállapítások jelentősen korlátozzák a trajektóriák, vagy a „világvonalak” fogalmának hasznosságát. Ekkor a méréseinkkel kompatibilis világvonalak együttesét kell vizsgálnunk. De mielőtt elhagynunk egy trajektóriát a vizsgálatból, elhagyjuk a szigorú determinizmus modelljét is. Csak statisztikai előrejelzéseket tudunk tenni, vagyis *előrejelzéseink átlagok lesznek*.

Ennek ellenére a klasszikus ortodoxia propagálói éveken keresztül próbálták megszabadítani a kvantummechanikát a statisztikai aspektusától. Albert Einstein sokat idézett mondása: „Az Isten nem játszik kockajátékot.” Most viszont azt látjuk, hogy amikor hosszú periódusokat vizsgálunk, maga a klasszikus dinamika igényli a statisztikai módszereket. Ennél is fontosabb azonban, hogy a klasszikus mechanika, amely talán a legkidolgozottabb az összes elméleti tudomány közt, ugyan zárt tudománynak tekinthető, de nagyon nem mindenható, azaz a valóságnak csak kis részét, s azt is csak bizonyos mértékig modellezi jól.

A klasszikus mechanikában a pontrészcskékből álló rendszer állapotát a q_1, q_2, \dots, q_s koordinátákkal és a p_1, p_2, \dots, p_s momentumokkal írjuk le. A rendszer energiáját ezekkel a változókkal kifejezve, a következő összefüggést kapjuk:

$$H = E_{kin}(p_1, p_2, \dots, p_s) + V_{pot}(q_1, q_2, \dots, q_s),$$

amelyben a jobb oldalon az első tag csak a momentumoktól (impulzusoktól) függ és a kinetikai energiát mutatja, a második pedig a koordináták függvénye és a potenciális energiát jelöli. Az energia ezekben a változóban kifejezve egy Hamilton-függvénnyel definiált, amely központi szerepet játszik a klasszikus dinamikában. Ezekben a vizsgálatokban most csak konzervatív autonóm rendszereket tekintünk, vagyis olyanokat, amelyekben H explicite nem függ az időtől. A rendszer mozgását a Hamilton-egyenletekből kapjuk, ami a klasszikus dinamika törvényei szerint a következőképpen fejezhető ki:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \quad \text{és} \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \quad (i = 1, 2, \dots, s).$$

Képzeljünk el most egy olyan $2s$ -dimenziós teret, amelynek pontjait a $q_1, q_2, \dots, q_s, p_1, p_2, \dots, p_s$ koordináták határozzák meg. Ezt a teret nevezzük *fázistérnek*. Mindegyik mechanikai állapothoz ennek a térnek egy P_i pontját rendeljük hozzá. Az induló P pont helyzete a t_0 időpontban, együtt a Hamilton-függvénnyel, tökéletesen meghatározza a rendszer mozgását, a rendszer időbeli evolúcióját (csak esetleg nem tudjuk azt kiszámolni).

Ha vesszük a $q_1, q_2, \dots, q_s, p_1, p_2, \dots, p_s$ koordináták egy tetszőleges f függ-

¹⁰Külön kell választani a valóságot és a modellt. A szokásos modellünkben a Naprendszer örök időre stabil. A valóságban ki tudja?

vényét, felhasználva a Hamilton-egyenleteket, megkapjuk az időbeli változását:

$$\begin{aligned}\frac{df}{dt} &= \sum_{i=1}^s \left[\frac{\partial f}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial f}{\partial p_i} \frac{dp_i}{dt} \right] = \\ &= \sum_{i=1}^s \left[\frac{\partial f}{\partial q_i} \frac{dH}{dt} - \frac{\partial f}{\partial p_i} \frac{dH}{dt} \right] \equiv [f, H]\end{aligned}$$

amelyben $[f, H]$ az f -nek H -val való Poisson-zárójele. Ezért a feltétel az f invarianciájára $[f, H] = 0$. Világos, hogy

$$[H, H] = \sum_{i=1}^s \left[\frac{\partial H}{\partial q_i} \frac{dH}{dt} - \frac{\partial H}{\partial p_i} \frac{dH}{dt} \right] = 0.$$

Ez a kifejezés egyszerűen az energiamegmaradást fejezi ki.

A klasszikus dinamika és a termodinamika összekapcsolásában követjük Gibbst és Einsteint, és bevezetjük a reprezentatív együttes fogalmát. Az alapötlet az, hogy egyetlen egy dinamikai rendszer helyett olyan rendszerek kollekcióját vizsgáljuk, amelyek mindegyike ugyanazon Hamilton-függvénynek feleltethető meg. Ennek a kollekciónak vagy együttesnek kiválasztása függ a rendszerekre tett feltevésektől és a kezdeti feltételektől. Amennyiben a kezdeti feltételek jól definiáltak, akkor az együttes biztosan a fázistér bizonyos régiójába koncentrálódik, míg ha kevésbé jól definiált, az együttes egy széles régió felett fog eloszlani a fázistérben. A rendszerek Gibbs-féle együttese egy pontfelhővel reprezentálható a fázistérben. Határértékben, amelyben mindegyik régió nagyszámú pontot tartalmaz, a felhő egy folytonos folyammal írható le a fázistérben, amelynek sűrűségét jelölje a

$$\rho(q_1, q_2, \dots, q_s, p_1, p_2, \dots, p_s, t)$$

sűrűségfüggvény. Mivel a pontok száma az együttesben tetszőleges, ezért ρ -t normalizáljuk, azaz

$$\int \dots \int \rho(q_1, q_2, \dots, q_s, p_1, p_2, \dots, p_s, t) dq_1, \dots, dp_s = 1.$$

Ezért a $\rho dq_1, \dots, dp_s$ egy reprezentatív pont megtalálásának valószínűségét mutatja a fázistér dq_1, \dots, dp_s tömegű elemében.

A sűrűség változása a fázistér mindegyik tömegelemében megfelel a határain átmenő folyamatok differenciájának. Említésre méltó tulajdonság, hogy a folyamat a fázistérben „összenyomhatatlan”. Más szavakkal, a folyamat divergenciája eltűnik. Valóban, a Hamilton-egyenletek felhasználásával azt kapjuk, hogy

$$\sum_{i=1}^s \left[\frac{\partial}{\partial q_i} \left(\frac{dq_i}{dt} \right) + \frac{\partial}{\partial p_i} \left(\frac{dp_i}{dt} \right) \right] = 0. \quad (3)$$

Az eredmény: a tömeg a fázistérben megőrződik időben. Felhasználva a fenti (3) egyenletet, egy egyszerű mozgási egyenletet kapunk a ρ sűrűségű fázistérre. Ez a jól ismert *Liouville-egyenlet*, amely a következő alakban írható

$$\frac{\partial \rho}{\partial t} = - \sum_{i=1}^s \left[\frac{\partial H}{\partial p_i} \frac{\partial \rho}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial \rho}{\partial p_i} \right] = [H, \rho], \quad (4)$$

s amelyben a $[\cdot, \cdot]$ zárójel most a H -nak ρ -val való Poisson-zárójele. Mivel gyakran kényelmesebb operátor megfogalmazást használni, egyszerűen megszorozva a (4) egyenletet $i = \sqrt{-1}$ -vel, írhatjuk

$$i \frac{\partial \rho}{\partial t} = L \rho,$$

amelyben L most egy lineáris operátort jelöl:

$$L = -i \frac{\partial H}{\partial p} \frac{\partial}{\partial q} + i \frac{\partial H}{\partial q} \frac{\partial}{\partial p}.$$

A klasszikus dinamikában az L a fázistérre, míg a kvantummechanikában a koordináta- vagy a momentum-térre vonatkozik. Az is jól látható, hogy a rendszer időbeli evolúcióját az L *Liouville-operátor* írja le az együttes közelítésben.

Az együttes-elmélet fontossága nyilvánvaló. Még ha nem is ismerjük a pontos kezdeti feltételeket, de vizsgálhatjuk a Gibbs-i sűrűséget és kiszámíthatjuk a tetszőleges mechanikai tulajdonság átlagos értékét, az $\langle A \rangle$ -t, úgy, hogy

$$\langle A \rangle = \int \int A(p, q) \rho dq dp,$$

amelyben felhasználtuk az együttes átlagot, $A(p, q)$ -t.

Megjegyezzük, hogy a Liouville-egyenlet formális megoldása könnyen előállítható a következő alakban:

$$\rho(t) = e^{iLt} \rho(0).$$

Állításunk könnyen igazolható a kifejezés közvetlen differenciálásával. További megjegyzést érdemel itt még, hogy a Gibbs-i együttesközelítés a valószínűség fogalmát a ρ sűrűségfüggvényen keresztül vezette be a fázistérbe. Ez lehetővé teszi mind a tiszta esetek tanulmányozását, amelyekre az induló feltételeket előírjuk, mind a kevert esetekét, amelyek különféle lehetséges kezdeti feltételeknek felelnek meg. Mindenesetre a sűrűségfüggvény időbeli evolúciója szigorúan determinisztikus. Nincs olyan egyszerű kapcsolata a valószínűségi folyamatokkal, mint a Brown-mozgás esetében van, és olyan fogalmak, mint átmeneti valószínűségek sem jelennek meg benne. Egy meglepő különbség az idő szerepében van: a Liouville-egyenlet formális megoldása érvényes az összes pozitív és negatív t értékekre.

Az operátorok fontos jellemzője a spektrumuk, amelyek egyszerűbb esetekben a sajátértékeiknek felelnek meg. Egy A operátornak a λ szám sajátértéke, ha

$$Au = \lambda u$$

teljesül az operátor értelmezési tartományában levő valamely nem nulla u esetén.

Az operátorok spektrumának szemléltetésére tekintsük a másodrendű differenciáloperátort valós függvényeken értelmezve; formálisan:

$$A = \frac{d^2}{dx^2}.$$

Ez a formula azonban önmagában nem ad egyértelmű meghatározást, ahhoz még meg kell adni azt a függvényosztályt, amelyen értelmezzük. A függvényosztályt a határfeltételekkel választjuk ki. Legyen például a határfeltétel az, hogy a függvények az $x = 0$ és $x = L$ helyeken nulla értéket vegyenek fel. Ekkor a sajátértékegyenlet megoldásai

$$\lambda = -k^2$$

és

$$u(x) = \sin kx,$$

ahol a határfeltétel $\sin kL = 0$ egyenlőségéből azt kapjuk, hogy

$$k = \frac{n\pi}{L}$$

tetszőleges $n \neq 0$ egész szám esetén. A sajátértékek függnek L -től, „végtelen nagy” intervallum esetén folytonos spektrumot kapunk.

Ha más határfeltételt adunk meg, például azt, hogy ne a függvény, hanem az első deriváltja vegyen fel nulla értéket az $x = 0$ és $x = L$ helyeken, akkor a sajátértékegyenlet megoldásai

$$u(x) = \cos kx,$$

ahol k -t az előbbi képlet adja meg, de most az $n = 0$ érték is megengedett.

4. A kvantummechanika dinamikája

A 20. század nagy fizikai elméletei a kvantummechanika, a speciális relativitáselmélet, az általános relativitáselmélet és az erőterek kvantumelmélete. Ezek szorosan függnek egymástól: az általános relativitáselmélet a speciálisra épül, az erőterek kvantumelmélete vagy a kvantum-térelmélet kiindulását pedig a speciális relativitáselmélet és a kvantummechanika alkotja. Ugyanakkor az általános relativitáselmélet határesetként tartalmazza a Newton-féle elméletet, a klasszikus mechanikát.

Ez az időszak volt a funkcionálanalízis kialakulásának időszaka is. A Lebesgue-integrál ráirányította a figyelmet a különféle függvényterekre, ekkor született meg az ortogonális sorfejtések és az integráloperátorok elmélete. Neumann János ekkor ismerkedett meg Göttingenben Heisenberg kvantummechanikájával és Hilbert integráloperátoraival. A kvantummechanika alapjait vizsgáló első cikkét 1927-ben írta Hilberttel és Nordheimmel együtt. Ugyanebben az évben jelenik meg „*A kvantummechanika matematikai megalapozása*” című dolgozata, amelyben először jelent meg a Hilbert-tér fogalma, amely azóta már klasszikussá lett: egy komplex vektortér, rajta értelmezett skaláris szorzattal, s azzal a tulajdonsággal, hogy a Cauchy-féle sorozatok konvergensek a skaláris szorzatból származtatott normára. Ez a nagyszerű elméleti konstrukció valósággal tálcán kínálja magát egy különös egyensúlyi dinamika aszimptotikusan stabil megközelítésére szinte valamennyi tudományterületen, köztük a közgazdaság-tudományban is.

A kvantummechanika volt az, amely megrázta a fizika Galilei által lefektetett alapjait. Eloszlatta azt a hitet, hogy a fizikai leírás realisztikus abban a naiv értelemben, hogy a fizika nyelvezete a kísérletezési és mérési feltételektől független rendszer tulajdonságait képviseli. Mikroszkopikus elmélet, minthogy az atomok és a molekulák viselkedését írja le. Benne a kvantumrendszer állapotát a Schrödinger hullámfüggvény határozza meg, amely egy időben reverzibilis dinamikus egyenletből származik, ugyanúgy, mint a klasszikus dinamikában a Hamilton-függvény. Bizonyos mérések azonban megzavarhatják a kvantumrendszereket. Ebből ered az a közkeletű tévedés, hogy a kvantummechanikában az irreverzibilitás oka a mérés. A folyamatok nyilván függetlenek attól, hogy figyeljük-e őket vagy sem, és azok a folyamatok is irreverzibilisek, amelyek megfigyelése semmiféle zavart sem okoz bennük.

Amíg a klasszikus mechanikában a Hamilton-függvény a koordináták és a momentumok függvénye, addig itt most a legegyszerűbb kvantum esetben is, például a hidrogénatom tulajdonságainak interpretálásában is egy Hamilton-operátorra lesz szükségünk, hogy a megfelelő energiaszinteket az operátorhoz kapcsolódó sajátértékekkel azonosíthassuk. A kérdéses sajátérték-egyenlet a következő:

$$H_{op}u_n = E_n u_n, \quad (5)$$

ahol az E_1, E_2, \dots, E_n sajátértékek a rendszer energiaszintjei. Természetesen megfelelő szabályokra van szükségünk, hogy a klasszikus változókról a kvantum operátorokra válthassunk, amelyek most a következők:

$$q \rightarrow q_{op} \quad \text{és} \quad p \rightarrow p_{op} = \frac{\hbar}{i} \frac{\partial}{\partial q},$$

azaz a koordináták nem változnak, a momentumokat pedig kicseréljük a koordináták szerinti deriváltakkal, ahol a transzformációban \hbar a Planck-állandót jelöli¹¹. A függvényekről az operátorokra történő áttérést a spektroszkopikus vizsgálatok

¹¹ Az utóbbi transzformáció részleteit lásd in Neumann (1980) 112–15. oldalakon.

tették szükségessé, amelyek kimutatták az energiaszintek létezését. Tulajdonképpen az operátorok bevezetése változtatta meg radikálisan a természet leírását. A rendszer állapotát a Hilbert-tér elemei adják, az operátorokat a Hilbert-térben értelmezzük. Az operátorok nem kommutatívak, ami most fizikai értelemben azt jelenti, hogy nem lehetséges olyan állapot, amelyben mind a q koordináták, mind a p momentumok egyidejűleg jól definiáltak. Ezt fogalmazza meg lényegében a *Heisenberg-féle határozatlansági reláció*, amely a kvantummechanika egyik alappillére. A két mennyiség határozatlanságának szorzata mindig nagyobb kell, hogy legyen, mint a Planck-állandó 2π -ed része, azaz $\Delta q \Delta p \geq h/2\pi$. Ha tehát egyre pontosabban szeretnénk megadni egy részecske helyét, akkor egyre bizonytalanabb lesz momentumának megadása és fordítva. Mindebből az következik, hogy a mikrofizika birodalmában gyökeresen más törvények léteznek, mint a klasszikus fizikában. Nincsenek pontosan meghatározható tények, csak valószínűségek vannak. Ez a felismerés szöges ellentétben állt Hegel történetfilozófiájával és később Marx determinisztikus gazdasági és társadalmi fejlődéstörvényeivel, de a fizikában is forradalmi következményekkel járt együtt.¹² Például az energia megmaradásának törvénye csak bizonyos feltételek mellett érvényes. Az energia és az idő is komplementer mennyiségpárok, egy adott pillanatban nem értelmezhető az idő és az energia közötti határozatlanság. Mindebből az következik, hogy az ilyen részecskékből álló rendszer összenergiája egyik pillanatról a másikra teljesen véletlenszerűen megváltozhat. Minél rövidebb a vizsgált időtartam, annál nagyobbak lehetnek ezek a fluktuációk, azaz az energiaingadozások.

Niels Bohr fogalmazta meg a *komplementaritási elvet*, amely a nem-kommutatív operátorok¹³ által képviselt fizikai mennyiségek egzisztenciáján alapul és azt mondja ki, hogy a kvantumos jelenségek körében a hullám és a részecske tulajdonságok egymást kiegészítik. Prigogine értelmezésében ez azt jelenti, hogy „a világ sokkal gazdagabb annál, mint ami bármely nyelven kifejezhető”.

A sajátfüggvények nagyon sokban ugyanazt a szerepet játsszák, mint a bázisvektorok a vektoralgebrában. Amint egy tetszőleges vektor az elemi matematikában kifejezhető a bázisvektorok bizonyos koordinátákkal alkotott lineáris kombinációjaként, úgy egy kvantummechanikai rendszer tetszőleges Ψ állapota (hullámfüggvénye) a megfelelő sajátfüggvények szuperpozíciójaként reprezentálható:

$$\Psi = \sum c_n u_n. \quad (6)$$

¹²Itt megjegyzést érdemel, hogy egy kérdéses kvantumelmélet mindaddig determinisztikus, ameddig ismerjük a hullám időbeli alakulásának törvényszerűségeit. Vagyis, ha a hullám jól definiált valamely időpontban, akkor kiszámíthatjuk, hogy milyen lesz valamely más időpontban. Az előrejelezhetetlen, véletlenszerű elem létezésére számos kísérlet utal, amit persze tovább bonyolít az a megfigyelés, hogy a kvantummechanikában az indeterminizmus (a valószínűség) időbeli változása determinisztikus.

¹³Itt érdemes megemlíteni, Neumann *operátorgyűrű* elméletét, amelyben a gyűrű nem-kommutatív operátorokból áll, vagy a mai nevén a *Neumann-algebrákat*. Ennek definiálásához szükségünk lesz az operátor halmaz kommutánsa fogalmára: azon X operátorok, amelyek a kérdéses halmaz minden egyes operátorával kommutálnak. Maga a Neumann-algebra pedig olyan operátor-algebra, amely megegyezik kommutánsának a kommutánsával.

Az egyszerűség kedvéért, itt érdemes venni a sajátfüggvények ortonormált halmazát (ami a vektoralgebrában annak felel meg, hogy a bázisvektorok egységnyi hosszúságúak és páronként ortogonálisak) :

$$\langle u_i | u_j \rangle = \delta_{ij} \begin{cases} = 1 & \text{ha } i = j \\ = 0 & \text{ha } i \neq j \end{cases}, \quad (7)$$

ahol az $\langle u_i | u_j \rangle$ jelölés egyféle skaláris szorzatot jelöl a következő művelettel definiálva:

$$\langle u_i | u_j \rangle = \int u_i^\dagger u_j dx,$$

amelyben az u_i^\dagger az u_i komplex konjugáltja, és az integrál az egész értelmezési tartományra kiterjesztendő. Megszorozva a (6) egyenletet u_i^\dagger -vel és felhasználva a (7) egyenletben adott ortonormalitási feltételeket, azonnal láthatjuk, hogy

$$c_n = \langle u_n | \Psi \rangle. \quad (8)$$

Az elemi vektortér és a kvantummechanikában használt tér közötti fő különbség az, hogy a dimenziószám az elsőben véges, a másodikban végtelen. A második esetben beszélünk Hilbert-térről, és az u_n sajátfüggvények, vagy a Ψ hullámfüggvények a tér elemei. Mindegyik elem kétféleképpen jelenhet meg a fizikai skaláris szorzaton belül, balra vagy jobbra. E célból vezetett be Dirac (1939) egy elegáns jelölést: az u_n elem írható mint egy *bra* vektor $\langle u_n |$, vagy mint egy *ket* vektor $| u_n \rangle$. Ezek a jelölések lehetővé teszik, hogy kompakt módon jelöljük a Hilbert-tér fontosabb tulajdonságait. Tegyük fel, hogy a (6) egyenletbeli kiterjesztés érvényes minden elemre. Felhasználva a *bra-ket* jelölést és a (8) egyenletet, egy tetszőleges Φ elemre írhatjuk:

$$| \Phi \rangle = \sum_n c_n | u_n \rangle = \sum_n | u_n \rangle \langle u_n | \Phi \rangle.$$

Mivel ennek a relációnak igaznak kell lennie bármely tetszőleges $| \Phi \rangle$ elemre, kapjuk a teljességi relációt

$$1 = \sum_n | u_n \rangle \langle u_n |.$$

A c_n expanziós koefficienseknek (azaz a kifejtési vagy Fourier-együtthatóknak), amelyek a (6) egyenletben jelennek meg, fontos fizikai jelentésük van. Ha mérjük a kérdéses fizikai mennyiséget (mondjuk az energiát), amelyből az u_n sajátfüggvény, az u_n -nek megfelelő sajátfüggvény megtalálásának valószínűsége $|c_n|^2$. A Ψ függvényt, amely a kvantum állapotát adja, ezért *valószínűségi amplitudónak* is nevezik. Vagyis a *kvantummechanika*, a klasszikus mechanikával szemben, egy *sztochasztikus-statisztikus elmélet*, amelyben a valószínűségi változók szerepét az

önadjungált operátorok (ld. később) veszik át, és bizonyos esetekben le kell mondanunk az együttes eloszlások használatáról.

A fentiekből világos, hogy a kvantummechanikában a fizikai mennyiségeket operátorok jelölik. Azonban ezek az operátorok nem lehetnek tetszőlegesek. A megfelelő operátorok specifikus osztálya az A operátornak az adjungáltja, amely a következő képlettel definiálható¹⁴:

$$\langle v | Au \rangle = \langle A^\dagger v | u \rangle. \quad (9)$$

Az önadjungált (vagy hermitikus) operátorok, amelyekre

$$A = A^\dagger,$$

fontos szerepet játszanak a kvantummechanikában. Fontosságukat az adja, hogy az önadjungált vagy hermitikus operátorok sajátértékei valósak.¹⁵ Továbbá, egy hermitikus operátor az (5)-öt kielégítő sajátfüggvények ortonormált halmazára vezet.

A hermitikus operátorokon túl szükségünk lesz még az operátorok egy másik osztályára is, amelyek a koordinátákbeli változásokkal kapcsolatosak. Az elemi geometriából ismert, hogy egy skaláris szorzat értékét bizonyos transzformációk nem változtatják meg. Ezért tekintsünk olyan A operátort, amely változatlanul hagyja a skaláris szorzatot. Ez azt implicálja, hogy

$$\langle Au | Av \rangle = \langle u | v \rangle$$

és felhasználva (9)-et, kapjuk

$$A^\dagger A = 1. \quad (10)$$

Per definitionem, a (10)-et kielégítő operátorokat *unitér operátoroknak* nevezzük. Az A operátor inverze, A^{-1} olyan, hogy

$$AA^{-1} = A^{-1}A = 1.$$

Ebből az is következik, hogy egy unitér operátor rendelkezik azzal a tulajdonsággal, hogy inverze megegyezik az adjungáltjával:

$$A^{-1} = A^\dagger.$$

¹⁴A definíció bizonyos pontatlanságot is tartalmaz, mert az értelmezési tartományok itt is fontos szerepet játszanak.

¹⁵Neumann János nevezte elsőként ezeket az operátorokat *maximálisan szimmetrikus* vagy *önadjungált operátoroknak*. Egy szimmetrikus operátorhoz több önadjungált operátor is tartozhat, de az is előfordulhat, hogy egy szimmetrikus operátornak egyáltalán nincs önadjungált kiterjesztése. Egyébként Neumann névéhez rendelhető a nemkorlátos lineáris operátorok elmélete is. Továbbá, Neumann János bizonyította be, hogy ha egy hermitikus operátor sajátértékei diszkrét, akkor a sajátfüggvények teljes ortogonális függvényrendszert alkotnak. Ezek szerint tehát bármely függvény sorbafejthető.

Csakúgy, mint az elemi geometriában, egy *hasonlósági transzformáció* gyakran végzendő az operátorokra is. Egy hasonlósági S az A -ból az \bar{A} -ba az alábbi reláción keresztül vezet:

$$\tilde{A} = S^{-1}AS.$$

Ezzel a következőképpen végezhetjük el a Hamilton-operátor diagonalizációját. Olyan Hamilton-operátort kell vennünk, amely tartalmazza a kinetikai hozzájárulást, a H_0 -t és a potenciális energiát, V -t. Ekkor kereshetjük a hasonlósági transzfomációt,

$$\tilde{H} = S^{-1}HS$$

egy unitér S operátorban kifejezve, amely az induló Hamilton-operátort diagonálisba transzmálja. Ez ekvivalens a sajátérték-egyenlet megoldásával az (5)-ben. Valóban, a H -t egy matrixként reprezentálhatjuk, és az (5) egyenlet mutatja, hogy a sajátfüggvényeit felhasználó reprezentációban a H egy diagonális mátrixszal reprezentálható:

$$\langle u_i | Hu_j \rangle = E_j \langle u_i | u_j \rangle = E_j \delta_{ij}.$$

Mint az előbbieken már láttuk, egy kvantumrendszer állapotát a Ψ állapotvektor, vagy hullámfüggvény írja le, míg az időváltozást leíró *Schrödinger-(hullám)egyenlet*:

$$i \frac{h}{2\pi} \frac{\partial \Psi}{\partial t} = H_{op} \Psi,$$

amelyben $i = \sqrt{-1}$ és h a Planck-konstans. Fontos megjegyezni, hogy ezt az egyenletet nem a kvantummechanikából vezette le Schrödinger, hanem csak feltette, ami kísérlettel igazolható. Látjuk, hogy az egyenlet parciális differenciálegyenlet, a koordináták szerinti deriváltak H_{op} -ban vannak, az idő szerinti derivált elsőrendű csakúgy, mint a Hamilton-egyenletekben. Mivel Ψ ismerete egy tetszőleges kezdeti t_0 időpontban, a megfelelő határfeltételekkel együtt, lehetővé teszi Ψ kiszámítását mind a jövőre, mind a múltra, ezért a determinisztikus szemlélet itt is érvényes, csakúgy, mint a klasszikus mechanikában. Az egyenletből könnyen kifejezhetjük a Ψ hullámfüggvényt:

$$\Psi(t) = e^{-iHt} \Psi(0).$$

Mivel H hermitikus, ezért mind a klasszikus, mind a kvantummechanikában a rendszer időbeli evolúcióját unitér transzformációban fejezzük ki. Ez viszont azt jelenti, hogy a Schrödinger-egyenletben a rendszer időbeli evolúciója csupán csak a koordinátákbeli változásnak felel meg.

5. Albert Einstein eredményei

A fentiek és az előző pont összegzéseként leszögezhetjük, hogy ezidáig még senki sem ért el nagyobb eredményeket az anyag statisztikai elméletében és specifikusabban a fluktuációk elméletében, mint Albert Einstein. A Boltzmann-formula

inverzén keresztül származtatta a makroszkópikus állapot valószínűségét a vele kapcsolatos entrópiában kifejezve. Ez a lépés döntőnek bizonyult a fluktuációk teljes makroszkópikus elmélete számára. Erre Einstein (1905) „*Brown-mozgás*” leírása volt a véletlenszerű folyamatok egyik legelső példája. A kémiai reakciók modellezése a Markov-láncokkal, ugyanezen gondolatvonal kiterjesztése. Végül, Einstein volt az első, aki felismerte a Planck-konstans általános értelmét, ami a hullám – részecske dualitásához vezetett. Einstein az elektromágneses sugárzást vizsgálta. Kb. húsz évvel később azonban, *de Broglie* kiterjesztette Einstein összefüggéseit az anyagra. Heisenberg, Schrödinger és mások foglalták ezeket az elgondolásokat matematikai keretekbe. De ha az anyag mind hullám, mind részecske, akkor a klasszikus determinizmus trajektória-felfogása érvényét veszti. Eredményként, a kvantummechanikával csak statisztikai előrejelzéseket végezhetünk. Einstein élete végéig ellenezte azt a felfogást, miszerint az ilyen statisztikai megfigyelések megfelelnek a természet objektív tulajdonságainak. Spinoza istenében, a természettel azonosított istenben, a mindenekfeletti (supréme) racionalitás istenében hitt. Ebben a koncepcióban nincs helye a szabad kreációnak, a kontingenciának, az emberi szabadságnak. Bármilyen kontingencia, bármilyen véletlenszerűség, ami úgy tűnik, hogy létezik, csak átmeneti. Ha úgy gondoljuk, hogy cselekedeteink szabadok, ez csak azért van, mert ignoráljuk azok igazi okait. Arra a kérdésre, hogy mi az irreverzibilitás, azt a választ adta, hogy az egy illúzió, egy szubjektív impresszió, ami kivételes kezdeti feltételekből származik. Paradox helyzet, hogy a kvantummechanikában végzett kutatásaiért Nobel-díjat kapott, pedig sohasem fogadta el, hogy a világegyetemet a véletlen igazgatja.

6. Termodinamikai egyensúly

A *termodinamika első törvénye*, egyszerű (homogén) rendszerekre

$$dE = dQ - pdV, \quad (11)$$

ahol E a belső energia, p a nyomás, V a térfogat, Q a hőátadás. Ez a formula azt fejezi ki, hogy a rendszer belső energiájának kicsiny dt idő alatti megváltozása egyenlő azzal a hővel, amelyet kapott a rendszer, plusz azzal a munkával, amelyet a rendszeren végeztek.¹⁶

Megfelelően lassú folyamatok esetén fennáll a

$$dS = \frac{dQ}{T} \quad (12)$$

egyenlőség, ahol S a rendszer entrópiája és T az abszolút hőmérséklete.

Kombinálva a (11) egyenletet a (12)-vel, az entrópia teljes differenciálját kapjuk az E és V változóiban:

$$dS = \frac{dE}{T} + p \frac{dV}{T}.$$

¹⁶Bővebb kifejtésért lásd Matolcsi (2004).

Gibbs általánosította ezt a formulát, hogy belefoglalja a komponensekbeli változásokat is. Jelöljük $n_1, n_2, \dots, n_\gamma$ -val a különböző komponensekben lévő mólok számát. Ekkor azt írhatjuk, hogy

$$\begin{aligned} dS &= \left(\frac{\partial S}{\partial E} \right) dE + \frac{\partial S}{\partial V} dV + \sum_{\gamma} \left(\frac{\partial S}{\partial n_{\gamma}} \right) dn_{\gamma} = \\ &= \frac{dE}{T} + \frac{p}{T} dV - \sum_{\gamma} \frac{\mu_{\gamma}}{T} dn_{\gamma}. \end{aligned} \quad (13)$$

A μ_{γ} mennyiségek per definitionem kémiai potenciálok, amelyeket Gibbs vezetett be, és ezért a (13) egyenletet az entrópia Gibbs-féle formulájának nevezik. A kémiai potenciálok maguk a termodinamikai változók függvényei, úgy, mint a hőmérsékleté, a nyomásé, a koncentrációé stb. Azok egy különösen egyszerű alakot vesznek fel az ún. ideális rendszerekre, amelyekben logaritmikusan függnek a mól-törtektől, az $N_{\gamma} = n_{\gamma} / (\sum n_{\gamma})$ -ektől:

$$\mu_{\gamma} = \zeta_{\gamma}(p, T) + RT \log N_{\gamma},$$

amelyben R a gázkonstans (egyenlő a Boltzmann-féle k konstans és az Avogadro-féle szám szorzatával), és a $\zeta_{\gamma}(p, T)$ a nyomás és a hőmérséklet bizonyos függvénye.

Az entrópiatermelés egyszerű értékelése válik lehetővé, ha feltesszük, hogy az entrópia az egyensúlyon kívül ugyanazon E , V és n_{γ} változóktól függ, mint az egyensúlyban. Megfelelő kémiai folyamatokra vonatkozó kiegészítő feltevések mellett (részletesen lásd in Prigogine (1980, 84–85. o.)) a (13) Gibbs-féle formulából az egyensúly bizonyos környezetében az alábbi összefüggést nyerjük:

$$\frac{dS}{dt} = \sum_j X_j J_j \geq 0, \quad (14)$$

amelyben dS/dt az egységnyi idő alatt termelt entrópiát, J_j az előforduló különböző irreverzibilis folyamatok (kémiai reakciók, hőáramlás, diffúzió stb.) rátáit, az X_j pedig a megfelelő általánosított erőket (affinitásokat, a hőmérsékleti és a kémiai potenciálok gradienseit stb.) mutatja. Ez az irreverzibilis folyamatok makroszkopikus termodinamikájának alapképlete.

A termodinamikai egyensúlyban a

$$J_j = 0, \quad X_j = 0$$

kikötések szimultán teljesülnek az entrópia termelésében előforduló összes irreverzibilis folyamatra. Az alapképletből világos, hogy egy (öko)gazdasági modellben a Gibbs-féle formula csak akkor használható fel, ha az a termodinamikában definiált potenciálokkal és azokhoz kapcsolódó irreverzibilis folyamatok rátáival izomorf módon értelmezett. Az ilyenféle megfeleltetés kérdését elsőként Neumann János (1945) vetette fel, és azóta számtalan kísérlet történt sejtésének igazolására, illetve cáfolatára. A két legérdekesebb tanulmány Bródy Andrásé (1989) és Paul

A. Samuelsoné (1992), amelyeket felhasználunk a következő pontban Neumann sejtésének cáfolatában.

7. Bródy versus Samuelson kísérletei

Mielőtt nagyító alá tennénk *Bródy* (1987) tanulmányát, megmutatjuk, hogy semmiféle maximum vagy minimum nem értelmes egyetlen homogén rendszerre. Szigorúan termodinamikai összefüggésekből indulunk ki, ami végül is *Samuelson* (1992) tanulmányában található (15a)–(15c) képletekben az entrópiamaximumról, illetve a (16a)–(16c) képletekben az energiaminimumról kifejtett érvelésekre vezet.

Vegyünk egy rendszert, amely két egymással kölcsönhatásban levő alrendszerből áll, és tegyük fel, hogy az alrendszerek $E_1 + E_2$ összenergiája és $V_1 + V_2$ összterfogata állandó; ekkor az

$$S_1(E_1, V_1) + S_2(E_2, V_2)$$

összentrópia maximális lesz egyensúlyban.

Vegyünk most azt a rendszert, amely az előző két alrendszerből áll, azzal a feltétellel, hogy az alrendszerek $S_1 + S_2$ összentrópiája és $V_1 + V_2$ összterfogata állandó; ekkor az

$$E_1(S_1, V_1) + E_2(S_2, V_2)$$

összenergia minimális lesz egyensúlyban.

Látható, hogy a két rendszer merőben különböző: az elsőben a két alrendszert együttesen hőszigeteljük és merev fallal vesszük körül, a másodikban a merev fal marad, de a hőszigetelés helyett éppen azt biztosítjuk, hogy mindig annyi hőt vezetünk ki vagy be, hogy az összentrópia állandó maradjon. A két rendszer kizárja egymást. *Nem lehet tehát ugyanarra a rendszerre egyszerre entrópiamaximumot és energia-minimumot számolni.* Az első rendszert jellemző mennyiségek E_1, E_2, V_1, V_2 , ezekre fogalmazunk meg feltételeket és ezek egy függvényének, az összentrópiának keressük a maximumát. A második rendszert más mennyiségek jellemzik, ezekre fogalmazunk meg feltételeket és ezek egy másik függvényének, az összenergia minimumát keressük. Összefoglalva: két különböző rendszerünk van, két összegeztethetetlen feltétellel, és két különböző függvény szélsőértékét keressük.

Bródy tanulmányában egyetlen rendszer van, a p és x mennyiségekre felírt két összegeztethető $pA \leq \lambda pB$ és $Ax \geq \lambda Bx$ Neumann-feltétellel, és ugyanannak a λ függvénynek (egyszer mint entrópiának, másszor mint energiának) keresi a szélsőértékét. Mindez a fentiek alapján kizárja a termodinamikával vett párhuzamot. Másképpen, *Neumann János* sejtése nem igazolható, hogy a gazdasági modelljében a termodinamikában definiált potenciálokkal és azokhoz kapcsolódó irreverzibilis folyamatok rátáival izomorf módon értelmezhető lenne a növekedés.

Viszont ez még nem zárja ki annak lehetőségét, hogy általánosabb keretek közt mégis bizonyítható.¹⁷

8. Stabilitás

A termodinamikai egyensúlynak megfelelő állapotok vagy egy minimális entrópiatermelésnek megfelelő stacionárius állapotok a lineáris nemegyensúlyi¹⁸ termodinamikában általában stabilak. Az entrópia a lineáris nemegyensúlyi termodinamikában Ljapunov-függvény a stabilitásra; az entrópiatermelés valójában a Ljapunov-függvény „rendszer szerinti driváltja”. Az entrópia második deriváltjának negatív definitésége biztosítja a Ljapunov-függvény szélsőértékét az egyensúlyban, az entrópiatermelés pozitivitása – lásd a (14) formulát – pedig a Ljapunov-függvény rendszer szerinti deriváltjának a szélsőértékét. Tehát az entrópia létezése biztosítja az összes fluktuáció csillapodását, a rendszer stabilitását.

Az entrópia második deriváltjának negatív definittségét azzal az esettel szemléltethetjük, amikor egyedül csak a belső energia változhat. Ugyanis az entrópiára érvényes összefüggések szerint

$$\frac{\partial S}{\partial E} = \frac{1}{T},$$

$$\frac{\partial^2 S}{\partial E^2} = -\frac{1}{T^2} \frac{\partial T}{\partial E} < 0,$$

minthogy

$$\frac{\partial T}{\partial E} = \frac{1}{C_V},$$

ahol C_V az állandó térfogaton vett fajhő, amely pozitív.

A lineáris nemegyensúlyi termodinamika mindenképpen homogén testekre vonatkozik. Kérdés, mennyiben lehet felhasználni eredményeit az inhomogén rendszerekre, vajon az egyensúlytól távolabbi rendszerekre extrapolálható-e ez a stabilitás? Vannak válaszok az irodalomban (pl. Glansdorff-Prigogine (1971)), de ezek még nem adnak teljes körű megoldásokat. Az egyensúlyi tartománytól messze távol a helyzet változik. Ott ugyanis a kémiai kinetika játszik lényeges szerepet, s a kémiai kinetika bizonyos típusaira a rendszer instabillá válhat. Ez azt mutatja, hogy jelentős különbség van az egyensúlyban levő rendszerekre vonatkozó törvények és az egyensúlytól távolabb lévő nemegyensúlyi rendszerekre vonatkozó törvények között. Az egyensúly törvényei univerzálisak, azonban távol az egyensúlytól a viselkedés nagyon specifikussá válhat.

¹⁷Vitáinkon ezt az álláspontot képviselte mindvégig Martinás Katalin és Ván Péter is.

¹⁸Érdemes itt megjegyezni, hogy a nemegyensúlyi értelmezés nem adható meg a szokásos (egyensúlyi) keretek között. Ha valahogy mégis értelmezhetjük a nemegyensúlyit, akkor az általános nemegyensúlyinak speciális esete az, amelyik közel van az egyensúlyhoz. Egyféle formális közelítése a fentieknek a híres *Onsager-elmélet*, amelyet Matoicsi (2004) kitűnően interpretál.

9. Önszerveződés: instabilitás, bifurkáció és evolúció

Elméletileg a *bifurkáció* egyszerűen az egyenletek új megoldásának megjelenése bizonyos kritikus érték(ek)re. Ezek a megoldások egymást követő *instabilitások* mellett nyerhetők, amelyek akkor fordulnak elő, amikor egyre jobban eltávolodunk az egyensúlytól. Például, vizsgáljunk meg egy kémiai reakciót, amely az alábbi rátaegyenletnek felel meg:

$$\frac{dX}{dt} = \alpha X (X - R).$$

Nilvánvaló, hogy az $R < 0$ -ra az egyetlen időfüggetlen megoldás az $X = 0$. Az $R = 0$ pontban egy új megoldás bifurkációját kapjuk, az $X = R$ -t. Könnyen belátható, hogy az $X = 0$ megoldás instabil, míg az $X = R$ stabil. Általában egymást követő bifurkációkat kapunk, ha növeljük a megfelelő α karakterisztikus paraméter értékét. A bifurkáció matematikai elmélete általában nagyon komplex. Gyakran nagyon fáradságos kifejtést igényel, de vannak esetek, amelyekben pontos megoldás lehetséges. Egy nagyon egyszerű ilyen típusú helyzetet ír le *René Thom* (1975) *katasztrófa elmélete*, amely akkor alkalmazható, amikor a diffúziót negligáljuk az X_i evolúciójából, és csak egy potenciálból származtatjuk. Ekkor a következő alakot írhatjuk fel:

$$\frac{dX_i}{dt} = -\frac{\partial V}{\partial X_i}, \quad (15)$$

amelyben V egy potenciálfüggvényt jelöl. Ez egy meglehetősen kivételes eset. Azonban, amikor teljesül, a (15) egyenlet megoldásainak egy általános klasszifikációja lehetővé teszi azon pontok megkeresését, amelyekben vannak változások a tartós állapotok stabilitási tulajdonságaiban. Ezek azok a pontok, amelyeket Thom a katasztrófák együttesének nevezett.

Végül egy általános fogalmat említünk meg, amely fontos szerepet játszik az *önszerveződés* elméletében, mégpedig a *strukturális stabilitást*¹⁹. A fogalom a legegyszerűbben a *ragadozó-áldozat versengésnek* megfelelő Lotka-Volterra egyenletek²⁰ egyszerűsített alakjával illusztrálható:

$$\begin{aligned} \frac{dx}{dt} &= by, \\ \frac{dy}{dt} &= -bx. \end{aligned} \quad (16)$$

A (16) rendszer (x, y) fázisterében a zárt trajektóriák végtelen halmaza veszi körbe az origót. A rendszer Jacobi-sajátértékei tisztán imaginárius komplex számok. Hasonlítsuk össze a (16) rendszer megoldásait azokkal, amelyeket a következő egyen-

¹⁹Konkrét vizsgálatra lásd a Móczár - Krisztin (2006) tanulmányt, amelyben a szerzők a keynesi elmélet Harrod-féle növekedési modelljének strukturális stabilitását vizsgálják.

²⁰A Lotka-Volterra rendszer közgazdasági adaptációját lásd in Goodwin (1967)).

letekből kapunk:

$$\begin{aligned}\frac{dx}{dt} &= by + ax. \\ \frac{dy}{dt} &= -bx + ay.\end{aligned}$$

Az utóbbi esetben, még a legkisebb a ($a < 0$) paraméterérték mellett is, az $x = 0$, $y = 0$ pont aszimptotikusan stabil, vagyis egy olyan egyensúlyi pont, amelyhez az összes trajektória konvergál a fázistérben. Ekkor $(0, 0)$ -ban a Jacobi-sajátérték negatív valós résszel rendelkezik. Per difinitionem, a (16) egyenletek által meghatározott rendszert *strukturálisan instabilnak* nevezzük, mivel a rendszer kis perturbációja megváltoztatja a fázisképét.

A strukturális stabilitás fogalma alkalmasnak látszik arra, hogy a legkompaktabb módon kifejezzük az innováció ideáját, új mechanizmus és új faj megjelenését, amelyek kezdetben hiányoztak a rendszerből.

A *Brüsszelerátor modell* tanulmányaink szempontjából különösen érdekes, mivel a megoldások dinamikai tulajdonságainak széles választéka tanulmányozható elégségesen távol az egyensúlytól: a határciklusok, a nem uniformizált tartós állapotok, a kémiai hullámok. Amikor a kémiai kinetikába a diffúziót is belefoglaljuk, a reakció-diffúzió egyenletek az alábbi alakot veszik fel a Brüsszelerátor modellben (részletesen lásd in Nicolis-Prigogine (1977)):

$$\begin{aligned}\frac{\partial X}{\partial t} &= A + X^2Y - BX - X + D_X \frac{\partial^2 X}{\partial r^2} \\ \frac{\partial Y}{\partial t} &= BX - X^2Y + D_Y \frac{\partial^2 Y}{\partial r^2}.\end{aligned}$$

Tegyük fel, hogy kezdeti értékeként a koncentrációs értékeket vesszük. A következő alakú megoldásokat keressük:

$$\begin{aligned}X &= A + X_0(t) \sin \frac{n\pi r}{L}, \\ Y &= \frac{B}{A} + Y_0(t) \sin \frac{n\pi r}{L},\end{aligned}\tag{17}$$

amelyben n egy integer szám és az X_0 és Y_0 még mindig időfüggők. Ezek a megoldások kielégítik a határfeltételeket, az $X = A$ -t és $Y = B/A$ -t az $r = 0$ -ra és $r = L$ -re. Ezután alkalmazhatjuk a lineáris stabilitáselemzést, amivel egy olyan diszperziós egyenletet nyerünk, amely a rendszerhez tartozó megfelelő Jacobi-sajátértéket hozzákapcsolja a (17)-beli n integer által adott tér-függőséghez.

Az eredmények a következők. Az instabilitás különbözőképpen állhat elő: a két diszperziós egyenletnek két olyan gyöke lehet, amelyek komplex konjugáltak, és valamely pontban a gyökök valós részei eltűnnek. Ez az eset határciklushoz vezet: az irodalomban ezt nevezik *Hopf-bifurkációnak*. Egy másik lehetőség, hogy két negatív valós gyökünk van, amelyek közül az egyik pozitívvá válik valamely kritikus pontban. Ez a helyzet vezet a térben nem uniformizált egyensúlyi állapotokhoz. Ezt nevezik *Turing-bifurkációnak*, mert Turing volt az első, aki megjegyezte

a morphogenezisről írt klasszikus tanulmányában (Turing (1952)), hogy egy ilyen bifurkáció lehetséges a kémiai kinetikában.

Lotka (1924) célkitűzése az volt, hogy a biológiai folyamatok egy új rendszerezését adja meg, és kifejtése tudatosan matematikai volt. Az egyszerű

$$dX/dt = aX$$

differenciálegyenleten alapuló növekedési folyamatok tanulmányozása után, Lotka bonyolultabb egyenleteket is vizsgált. Észrevette, hogy az ilyen folyamatokra az általános alak a

$$dX/dt = a + bX + cX^2 + \dots,$$

amely hamarosan olyan kísérletekre vezette, amelyek ilyen típusú egyszerű modellekre vezetnek. Egyik ilyen felfedezése volt a „populáció szaporodási törvénye”, amelyet a

$$dX/dt = aX + bX^2$$

differenciálegyenlettel definiált, az

$$X(t) = (a/b) / [1 + \exp(-at)]$$

megoldással, amely a közismert logisztikus görbe függvényalakja. Lotka hatalmas energiát szentelt arra, hogy kiderítse törvényének igazságát, vagy hamisságát. Ehhez különböző adatbázisokat, mint például az USA lakossága, a muslincák vagy a baktériumtelepek száma stb., használt fel statisztikai vizsgálataiban. Elképzelése helyes volt, mivel a differenciálegyenlete azon alapult, hogy egy populáció növekedésének pozitívan kell kapcsolódnia a populáció szintjével, és a növekedés megáll, amikor eléri az abszolút limitet.

Lotka az evolúciót, mint „az irreverzibilis változásokon átmenő rendszer múltját” definiálja. Olyan evolúciós törvényt célzott meg, amely pontosan úgy funkcionál, mégpedig az általánosság ugyanolyan fokával, mint a termodinamika törvényei, különösen a termodinamika második törvénye. A törvény egyetlen irányt állapított meg az izolált rendszerben megjelenő folyamatokra, és ezért egy *evolúciós törvénynek* tekinthető.

A közgazdasági dinamikában a számunkra érdekes legfontosabb eszme azonban az, hogy az alapul szolgáló modell, amelyen belül a keretet differenciálegyenletek szolgáltatják, a *disszipatív dinamikus rendszerek*hez tartoznak. A szokásos dinamikus rendszerek *konzervatívak*, amelyekben bizonyos mennyiség, általában az energia, a rendszer pályái és mozgásai mentén konzervált. Lotka kimutatja, hogy ezek a rendszerek reverzibilisek, vagyis az idő visszafordítható bennük, ami analitikus megoldásokat nyújt a mozgások korábbi helyzeteire. A keringő bolygók mozgását leíró fizikai rendszer megoldható előre, hogy megállapítsuk a Mars jövőbeli helyzetét, vagy visszafelé, hogy megállapítsuk a korábbi helyzetét. Lotka, mint a fejlődést kutató biológus, olyan matematikai modell-osztályok kifejlesztésében volt érdekelt, amelyek irreverzibilisek, vagyis amelyek nem fordíthatók visszafelé. Egy ilyen rendszerre az egyensúly fogalma nem a Naprendszernek, az égi mechanikának megfelelő periodikus mozgás volt, sem a hasonló fizikai egyensúlyoké, hanem inkább egy

disszipatív rendszerre vonatkozó egyensúly, egyféle sűrűlódásos Naprendszer, hővesztéssel járó, irány nélküli időfolyam. Biológiai egyedek kipusztulnak, szervezetek fejlődnek. Lotka szerint a társadalmi tudományoknak is ilyen *disszipatív rendszereket* (amelyeket Samuelson (1947) elsőfajú stabilitással jellemez Picardra (1928) hivatkozva) kell modelleznie. Lotka a nyílt és zárt rendszerek közötti distinkciójával és a velük kapcsolatos különböző egyensúlyi fogalmakkal bizonyos mértékben az általános rendszerelmélet (*von Bertalanffy* (1973)) lámpásának is tekinthető.

Lotka könyvének második része a kinetikát tárgyalja, és elsősorban a növekedés egyszerű egyenletét, az eredményül kapott logisztikus görbét vizsgálja. Az egyes fejezetek különböző populációkat és a növekedési törvény különböző példáit teszik nagyító alá. Samuelson természetesen átveszi ezt a kifejtést a *Foundations* 291–94. oldalain és hivatkozik Lotkára. Samuelson populáció elméletei hasonlóak Lotka populáció-növekedésről kifejtett elméleteivel.

A harmadik rész a statikáról szól, amelyben Lotkát az egyensúly érdekelte, vagyis az $F_i(X_1, X_2, \dots, X_n)$ függvények zérus értékekkel. Megjegyzi, hogy a kinetika szempontjából az egyensúly (vagyis egy stacionárius állapot) olyan állapot, amelyben a sebesség, a dX_i/dt zérussal egyenlő. Másképpen, az egyensúly dinamikus fogalma az erők mérlegéhez kapcsolódik, és megjegyzi, hogy ennek jelentése etimológiailag szorosan kapcsolódik az „egyensúly” mint „*aequa libra*” kifejezéshez. Ezeknél is érdekesebb a harmadik jelentés, az energetikában kifejezve, miszerint egy rendszer dinamikus egyensúlyban van, amennyiben az az energia dimenzióiban kifejezett bizonyos függvények minimumát (vagy néha a maximumát) elérte; egy olyan állapot, amelyben a virtuális munkavégzés tetszőleges kis elmozdítása kompatibilis a korlátok eltűnésével (i.m. p. 144). Az egyensúly változásaiban a

$$dX_i/dt = F_i(X_1, X_2, \dots, X_n; P)$$

alakú egyenleteket vizsgálta, amelyben a P lassan nő, úgyhogy a megoldások, vagy egyensúlyi állapotok, amelyeket a $P = \text{konstans}$ mellett származtattunk, folytonosan ahhoz az esethez kapcsolódnak, amelyben P nem egy konstans, hanem inkább egy lassú változó. Ezt követi a *Le Chatelier-elv* (a legkisebb kényszer elve), amely a következőképpen jellemezhető: „Mindegyik kémiai egyensúlyban levő rendszer, valamely egyensúlyi tényező változásának hatására, egy olyan irányú transzformáción megy keresztül, hogy ha kizárólag csak ez a transzformáció következik be, akkor a kérdéses tényező ellenkező irányú változását fogja eredményezni. Az egyensúlyi tényezők a hőmérséklet, a nyomás és az elektromos erő, amelyek megfelelnek a három energiaformának, a hőnek, az elektromosságnak és a mechanikai energiának.” (Lotka, 1924, 281. o). Másképpen, ha valamely rendszert (anyagot vagy anyagok halmazát) egyensúlyi állapotában megzavarnak, a rendszer úgy állítja helyre az egyensúlyát, hogy a zavaró hatást semlegesíti. Samuelson volt az, aki felhívta a közgazdászok figyelmét a *Le Chatelier-elvre*, és felhasználta a költségek és a termelés magyarázatában.²¹ Érdekes lehet itt megemlíteni, hogy Lotka példáiból kitűnik, hogy maga is tisztában volt korának matematikai közgazdasági

²¹Részletes kifejtését lásd in Zalai (2000) 442–445. o.

irodalmával, könyvében például hivatkozik is Cournotra, Edgeworthra, Jevonsra és Pareto-ra. Az X_i -k – könyvének negyedik részében – élő organizmusok aggregátumai; mint állapotváltozók valójában „energia transzformerek”, és a dinamika ezeknek a „motoroknak” a dinamikája. Ezért a koncepciójáért érdemelte ki később Lotka az „ökológia atyja” címet. Megjegyzi, hogy eszméi fényt deríthetnek kvantitatíve a közgazdaságtan biológiai alapjaira, azokra az összefüggésekre, amelyek bizonyos biológiai és bizonyos közgazdasági mennyiségek között léteznek. Az életért folyó verseny elsősorban az elérhető energiáért folyó küzdelem.²² Ebben az értelemben az energia értékes az organizmus számára, de ami nagyon különböző dolog attól, hogy azt mondjuk, hogy a közgazdasági érték az energia formája. A társadalmi fejlődés és a gazdasági szervezet energetikai közelítései Friedrich Wilhelm Ostwald elméleteit tükrözik, aki azt hirdette, hogy az energia a centrális szervező fogalom a fizikai és a biológiai tudományokban.

10. Időoperátorok

Prigogine (1980) a változások leírására alkalmas módszerek három osztályát különbözteti meg: (1) az átlagok evolúciójával foglalkozó makroszkopikus módszerek, mint pl. Fourier törvénye, a kémiai kinetika; (2) sztochasztikus módszerek, mint a Markov-láncok; (3) a klasszikus vagy kvantummechanika.

Az utóbbi években a makroszkopikus leírás területén valami egészen új és váratlan dolgok történtek, különösen a nemlineáris, az egyensúlytól távoli helyzetekben: még egyszerű példák is egymást követő bifurkációkhoz, különböző téridőbeli strukturákhoz vezethetnek. Ez viszont drasztikusan korlátozza a makroszkopikus leírás egységesítő erejét, és azt mutatja, hogy önmagában nem szolgáltathatja az időbeli evolúció konzisztens leírását. Ráadásul, a makroszkopikus egyenletek nem tartalmaznak információt arról, hogy mi történik a bifurkációs pontokban? Milyen lesz a rendszer törése adott bifurkációkat követően?

Ezért olyan sztochasztikus elmélet felé kell fordulni, mint a Markov-láncok. De itt is új dolgok jelennek meg. Különösen érdekes a fluktuációk és a bifurkációk közötti zárt összefüggés, ami mély alterációkhoz vezet a valószínűségelmélet klasszikus eredményeiben. A nagy számok törvénye többé nem érvényes a bifurkációk közelében, és a valószínűségi eloszlásokra vonatkozó lineáris mesteregyenletek megoldásának unicitása is elvész.

A sztochasztikus és makroszkopikus módszerek közötti összefüggés világos. Pontosan akkor észlelhető, amikor az átlagmennyiségek nem elégítik ki a zárt egyenleteket, amelyek a bifurkációs pontokhoz oly közel vannak, hogy a statisztikai elmélet teljes apparátusát kell használnunk. A makroszkopikus vagy sztochasztikus és a dinamikus módszerek közötti összefüggés azonban, továbbra is kihívás marad. Ezt a kérdést sokféle szempontból vizsgálták a múltban.

²²Lásd Hotelling (1931).

A kvantumelmélet is bizonyára szerepet játszik ebben, mert a klasszikus trajektoriák feladására kényszerít minket. De a második törvénnyel való kapcsolata szempontjából, az instabilitás fogalma, amelyet ismétlődően megvitattak, alapvető jelentőségűnek tűnik. A mozgási egyenletek strukturája a mikroszkopikus szinten levő „véletlenszerűséggel” ekkor a makroszkopikus szinten levő irreverzibilitásként jelenik meg. Ebben az értelemben már *Poincaré* (1921) előrejelezte az irreverzibilitás jelentését. Poincarénak az alapvetően determinisztikus leírásba vetett hite azonban túlságosan is szilárd volt, hogy a természet releváns statisztikai leírását komolyan vizsgálja. A helyzet mára már teljesen különböző. Sok évvel Poincaré után a természet determinisztikus leírásába vetett bizalmunk megingott, mind mikroszkopikus, mind makroszkopikus szinten. A klasszikus vagy kvantumfizika természetes korlátjai csökkenthetik az előrejelző erejüket. Bár az újonnan megjelent fogalmak segítenek a folyamat megfordításában. Ezek közül a legérdekesebb az M mikroszkopikus időoperátor. Az időoperátor pontos matematikai értelmezésére nagyon sok próbálkozás történt. Ezek sikere több szempontból is vitatott, s elsősorban az értelmezési tartományok definiálásán buknak el.

Az egyik ilyen próbálkozás Prigogine nevéhez fűződik, aki bevezeti a *második idő* fogalmát, amely egy belső idő és teljesen különbözik attól, amely megcímkézi a trajektoriákat vagy hullámfüggvényeket. Ez az időoperátor egy új bizonytalansági relációt teljesít az L Liouville-operátorral. A $\langle T \rangle$ és a $\langle T^2 \rangle$ átlagokat az alábbi bilineáris formulákon keresztül definiálhatjuk²³:

$$\langle T \rangle = \text{tr} \rho^\dagger T \rho, \quad \langle T^2 \rangle = \text{tr} \rho^\dagger T^2 \rho.$$

Eléggé érdekesen, a 'közönséges' idő, a dinamika címkéje, ekkor egy átlag lesz az új időoperátor felett. Ez valójában a bizonytalansági reláció következménye (vö. Prigogine, 1980, 188. o. és 209. o.), ami azt implikálja, hogy

$$\begin{aligned} \frac{d}{dt} \langle T \rangle &= \frac{d}{dt} \text{tr} \left[(e^{-iLt} \rho)^\dagger T (e^{-iLt} \rho) \right] = \\ &= i \text{tr} \left[\rho^\dagger e^{iLt} (LT - TL) e^{-iLt} \rho \right] = \\ &= \text{tr} \rho^\dagger \rho = \text{konstans}. \end{aligned}$$

Egy megfelelő normalizációval ezt a konstans egynek vehetjük. Így azt kapjuk, hogy

$$dt = d \langle T \rangle.$$

Szavakban, a makroszkopikus idő egyszerűen az új időoperátor feletti átlag. Ebben a perspektívában a szokásos idő fogalmát csak akkor nyerhetjük vissza, amikor T egy olyan triviális operátor (mint a klasszikus mechanikában), hogy

$$T \rho(x, v, t) = t \rho(x, v, t).$$

²³A klasszikus mechanikában az átlagolási operátor a fázistér feletti integrációt jelenti, amihez hasonló szerepet játszik a nyom operátor a kvantummechanikában, azaz $\text{tr} O = \sum_n \langle n | O | n \rangle$, és a sűrűség operátor, azaz $\rho = |\Psi\rangle\langle\Psi|$.

Ekkor a „kor” független az eloszlás alakjától a fázistérben. Ezzel szemben, az új fogalom azt implikálja, hogy a „kor” függ magától az eloszlástól, és többé nem egy külső paraméter, nem egy egyszerű címke, mint a konvencionális megfogalmazásban.

A Prigogine-féle időoperátor azonban a jelenlegi formájában hasonlóképpen értelmezhetetlen, mint a kvantummechanikai (lásd később a (b) pontban). Nem adja meg pontosan azt a függényszót, amelyen az operátor hat.

A kvantummechanikában az időoperátor pontos matematikai értelmezésére tett próbálkozások közül, most kettőt emelünk ki.²⁴

(a) Legyen X a térben négyzetesen integrálható komplex értékű függvények összessége, ami egy tömegpont állapotainak Hilbert-terét képezi. Az önadjungált

$$H = -\frac{\Delta}{2m} + V$$

Hamilton-operátor értelmezési tartománya az X -nek egy sűrű altere. Egy folyamat olyan függvény, amely minden t pillanathoz hozzárendel egy Ψ_t elemet X -ben; a folyamat kielégíti az

$$i\frac{d\Psi_t}{dt} = H\Psi_t$$

egyenletet. Minden fizikai mennyiség az X -ben (természetesen sűrűn) értelmezett önadjungált operátor. Így a T időoperátornak is ilyennek kell lennie, hogy a $HT - TH = i$ teljesüljön egy sűrű altéren (*kanonikus felcserélés*). Ilyen viszont nem létezik, mivel a t -vel való szorzás nincs értelmezve X -ben (mi az a t , amivel meg kellene szorozni egy a térben értelmezett függvényt?).

(b) Legyen C az időben és térben értelmezett komplex értékű kétszer differenciálható függvények összessége. Ezen értelmezett az idő szerinti differenciálás, valamint a

$$-\frac{\Delta}{2m} + V$$

másodrendű térbeli differenciáloperátor. Így értelmezhető Z_0 , mint azoknak a C -beli ϕ függvényeknek az összessége, amelyek kielégítik az

$$\left(-i\frac{\partial}{\partial t} - \frac{\Delta}{2m} + V\right)\phi = 0$$

egyenletet. Ezen térbeli integrálással skalárszorzatot értelmezünk, azzal teljessé téve Z_0 -t, kapjuk Z -t, s most ez képezi a tömegpont állapotainak Hilbert-terét. Ebben a felfogásban egy állapot maga egy folyamat. Minden fizikai mennyiség a Z -ben (természetesen sűrűn) értelmezett önadjungált operátor. Az az érdekesség, hogy itt az energia-operátor $i(\partial/\partial t)$. Formálisan ez és a t -vel való szorzás operátora kielégíti a kanonikus felcserélési relációt. Csakhogy a t -vel való szorzás csupán a C -ben értelmes, Z -ben nem. Ugyanis, ha ϕ a Z_0 eleme, akkor $t\phi$ már

²⁴Ezek megfogalmazásáért Matolcsi Tamásnak tartozom köszönettel.

nem az. És másképpen sem értelmezhető olyan önadjungált operátor, amely az energiával kanonikusan felcserél.

További próbálkozások abból indulnak ki, hogy általánosabb értelmezésben fizikai mennyiségnek nem kell önadjungált operátornak lennie, csak maximális szimmetrikusnak. Ilyen időoperátor viszont nincs.

11. Következtetések

Hosszú ideig a klasszikus mechanika abszolút előrejelezhetőségét úgy tekintették, mint a fizikai világ tudományos képének egy lényeges eleme. A modern tudomány három évszázada alatt (1685-től számítva, amikor is Newton előadta a *Principia* c. művét a Royal Society-ben) a tudományos világkép egy új, sokkal árnyaltabb fogalom irányába tolódott át, amelyben mind a determinisztikus, mind a sztochasztikus megközelítések jelentős szerepet játszanak. A sztochasztikus elemek nemcsak a makroszkopikus szint bifurkációs elméleteiben lényegesek, hanem a mikroszkopikus leírásban is, amint azt már a klasszikus mechanika is bebizonyította. Ez az evolúció még napjainkban is tart. Az is jól érzékelhető, hogy a *termodinamika törvényei nem kerülhetők meg az ökögazdasági modellekben*. Szigorúan a termodinamika törvényeiből kiindulva, Bródy (1989) és Samuelson (1992) tanulmányaira támaszkodva, megmutattuk, hogy Neumann János sejtése nem igazolható: a gazdasági modelljében a termodinamikában definiált potenciálokkal és azokhoz kapcsolódó irreverzibilis folyamatok rátáival nem értelmezhető izomorf módon a növekedés. Vagyis várat még magára egy olyan ökögazdasági modell, amely a termodinamika törvényein a valóság jobb közelítésére alkalmas. Az új időfogalom még nem jutott el a közgazdászokig, de az időoperátorok pontos értelmezése új távlatokat nyithat, különösen a makroökonómiai elméletekben. Erre talán, a megfelelő pontosítások után, leginkább Prigogine *második idő* fogalma aspirálhat. Az eddigi statikus egyensúlyi referencia pontokat felváltják a dinamikus, időben változó *sztochasztikus egyensúlyi referencia függvények*, ami forradalmian új megvilágításba helyez számos társadalomtudományi, s főleg közgazdasági kérdést. Mindez a dinamikai rendszerek új vizsgálatát eredményezi, új fogalmak és definíciók jelennek meg, új elméletek váltják fel az eddigi stabilitás-, bifurkáció-, evolúció- stb. elméleteket.

Hivatkozások

- [1] BARUCHA-REID, A. T.: *Elements of the theory of Markov processes and their applications*, New-York, McGraw-Hill, (1960).
- [2] BERTALANFFY, L.: *General system theory*. New-York: George Braziller, (1968).
- [3] BRÓDY ANDRÁS: „*Economics and Thermodynamics*”, in John von Neumann and Modern Economics, ed. by M. Dore, S. Chakravarty and R. Goodwin, Oxford, Clarendon Press, (1989).
- [4] CHANDRASEKHAR, S.: *Stochastic Problems in Physics and Astronomy*, Reviews of Modern Physics, 15(1), (1943), 1–89.

- [5] DEBREU, G.: *Excess Demand Functions*, Journal of Mathematical Economics, **1**, (1972), 15–23.
- [6] DIRAC, P. A. M.: *A New Notation for Quantum Mechanics*. Proceedings of the Cambridge Philosophical Society, **35**:416., (1939).
- [7] DIRAC, P. A. M.: *The principles of quantum mechanics*, 4th ed. Oxford: Clarendon, (1958).
- [8] EINSTEIN, A.: *On the movement of small particles suspended in stationary liquid demanded by the molecular-kinetic theory of heat*, Ann.d. Physik **17**, (1905) (In Investigations of the theory of Brownian movement, ed. R. Fürth, Dover, New York, 1956)
- [9] EINSTEIN, A., LORENTZ, H. A., WEYL, H., ÉS MINKOWSKI, H.: *The principle of relativity*, London: Methuen (Dover ed.), (1923).
- [10] GEORGESCU-ROEGEN, N.: *The Entropy Law and the Economic Process*, University of Alabama, Distinguished Lecture Series No.1, (1971).
- [11] GIBBS, J. W.: *Elementary Principles in Statistical Mechanics*, Yale University Press, Reprinted (1981) by Ox Bow Press, Woodbridge, CO, (1902).
- [12] GLANSDORFF, P. ÉS PRIGOGINE, I.: *Thermodynamic theory of structure, stability, and fluctuations*, New-York, Wiley-Interscience, (1971).
- [13] GOODWIN, R. M.: „A growth cycle” in Feinstein, C. H. (ed.): *Socialism, Capitalism and Economic Growth*, Cambridge, Cambridge University Press, (1967).
- [14] HAWKING, S. W.: *Az idő rövid története*, Budapest, Talantum, (1998).
- [15] HAWKING, S. W. – R. PENROSE: *A tér és az idő természete*, Budapest, Talantum, (1999).
- [16] HIRSCH: *Dinamikai rendszerek és differenciálegyenletek*, Alkalmazott Matematikai Lapok, **14**, (1989), 171–232.
- [17] HOTELLING, H.: *The Economics of Exhaustible Resources*, Journal of Political Economy, Vol. **39**, (1931), 137–175.
- [18] KOLMOGOROV, A. N.: *On the conservation of conditionally periodic motions under small perturbation of the Hamiltonian*. Dokl. Akad. Nauk. SSR **98**, (1954), 527–530.
- [19] KORNAI JÁNOS: *Antiequilibrium*, Közgazdasági és Jogi Könyvkiadó, Budapest, (1971).
- [20] LAGRANGE, J. L.: *Théorie des fonctions analytiques*, Paris: Imprimerie de la République, (1976).
- [21] LOTKA, A. J.: *Elements of Mathematical Biology*, reprinted by Dover, 1956, New York, (1924).
- [22] LJAPUNOV: *The general problem of the stability of motion*, Annales de la Faculté des Sciences de l'Université de Toulouse **9**(2):27:474, (1908).
- [23] MANTEL, R. R.: *On the characterizations of aggregate excess demand*, Journal of Economic Theory, **7**, (1974), 348–353.
- [24] MARTINÁS KATALIN: *About Irreversibility in Economics*, Open Sys. & Information Dyn. **7**, (2000), 349–364.

- [25] MATOLCSI TAMÁS: *Ordinary Thermodynamics (Nonequilibrium homogeneous processes)*, Akadémia Kiadó, Budapest, (2004).
- [26] MÓCZÁR JÓZSEF, KRISZTIN TIBOR: *A Harrod-modell strukturális stabilitása*, Szigma, XXXVII. évf., 1-2. szám (2006), 1-31.
- [27] NEUMANN JÁNOS: „*A Model of General Economic Equilibrium.*” Translated into English by C. Morgenstern. *Rev. Econ. Studies* 13:1-9 [VI, 3]), (1945).
- [28] NEUMANN JÁNOS: *A kvantummechanika matematikai alapjai*, Akadémiai Kiadó, Budapest, (1980/1932).
- [29] NICOLIS, G. ÉS PRIGOGINE, I.: *Self-organization in non-equilibrium systems*, New-York: Wiley, (1977)
- [30] PICARD. É.: *Traite d'analyse*, Paris: Gauthier-Villars, (1928).
- [31] PLANCK, M.: *Vorlesungen über Thermodynamik*, Leipzig, (1930).
- [32] POINCARÉ, H.: *Science and Hypothesis*, in *The foundations of science*, New-York: The Science Press, (1921).
- [33] PRIGOGINE, I.: *From Being to Becoming*, San Francisco, Freeman, (1980).
- [34] THOM, R.: *Structural stability and morphogenesis*. Reading, Massachusetts: Benjamin, (1975)
- [35] TURING, A. M.: *The chemical basis of morphogenesis*. *Phil. Trans. Roy. Soc. London*, Ser. B. 237:37, (1952), 37-72.
- [36] SAMUELSON, P. A.: *Foundations of Economic Analysis*, Harvard University Press, Cambridge, Mass, (1947).
- [37] SAMUELSON, P. A.: „*Economics and Thermodynamics: von Neumann's Problematic Conjecture*”, in *Rational Interaction*, ed. by R. Selten, Berlin, Springer-Verlag, (1992).
- [38] SHAFER, W., SONNENSCHN, H.: „*Market demand and excess demand functions*”, in *Handbook of Mathematical Economics*, Vol. 2, eds. K. J. Arrow - M. D. Intriligator, Amsterdam, North-Holland, (1982), 671-93.
- [39] SONNENSCHN, H. F.: *Market Excess Demand Functions*, *Econometrica*, Vol. 40, (1972), 549-563.
- [40] SONNENSCHN, H. F.: *Do Walras' Identity and Continuity Characterize the Class of Community Excess Demand Functions?*, *Journal of Economic Theory*, Vol. 6, No. 4, (1973), 345-54.
- [41] VÁN PÉTER: *Bevezetés a nemegyensúlyi termodinamikába*, Egyetemi jegyzet, BMGE, Budapest, (2005)
- [42] ZALAI ERNŐ: *Matematikai közgazdaságtan*, Közgazdasági és Jogi Könyvkiadó, Budapest, (2000)

(Beérkezett: 2009. július 27.)

MÓCZÁR JÓZSEF

Budapesti Corvinus Egyetem

Matematikai közgazdaságtan és Gazdaságelemzések Tanszék

jozsef.moczar@uni-corvinus.hu

THE NEWEST RESULTS OF PHYSICAL MATHEMATICS
AS THE POSSIBLE INVESTIGATION TOOLS OF ECONOMICS

JÓZSEF MÓCZÁR

The aim of this paper is to outline the newest results of physics, i.e., the stochastic mathematical relations of relativity theory and quantum mechanics as well as irreversible dynamics which can be applied for some economic problems. For example, the correct interpretation of time operators using for the macroeconomic theories may provide a serious improvement in approach to the reality. The stochastic dynamic equilibrium reference functions will take over the role of recent static equilibrium reference points, which may also reveal some nonequilibrium questions of macroeconomics. The concepts and definitions of thermodynamics and biological evolution have been adopted in economics by Paul A. Samuelson, but he did not concern the newest results of quantum mechanics, e.g., the time operators. Now we do it. In addition, following Samuelson, we show that von Neumann growth model cannot be explained as a peculiar extension of thermodynamic irreversibility.

KLAFSZKY EMIL

(1934–2009)



Klafszy Emil, a magyar operációkutatás fontos és sokak által szeretett alakja 2009. január 31-én elhunyt. Munkásságát a 2005. évi 1. számban ismertettük abból az alkalomból, hogy kitüntették az Egerváry Jenő Emlékplakettel. Most publikációs jegyzékét adjuk közre témák szerinti csoportosításban.

Publikációi

Első jelentős munkája

- *Hálózati folyamatok*, Bolyai János Matematikai Társulat, Budapest (1969), 263 o.

Kandidátusi értekezés

- Geometriai programozás és néhány alkalmazása (Kandidátusi értekezés), MTA SZTAKI, Tanulmányok **8.** (1973), 139 o.

Geometriai programozással kapcsolatos munkák

- *Geometriai programozás*, MTA Számítástechnikai Központ, Közlemények **8**, (1972) 41–65.
- *Marginális értékek a geometriai programozásban*, MTA Számítástechnikai Központ, Közlemények **9**, (1972) 51–68.
- *Geometric Programming*, IIASA Systems Analysis and Related Topics No. **11** (1976)
- *A lineáris cseremodell egyensúlyi árának meghatározása geometriai programozással*, *Alkalmazott Matematikai Lapok* **7**, (1981) 139–157.

Mérnöki tevékenységgel kapcsolatos munkák

- *Thermal Steady State Analysis* (Ottmár Bélával és Pattantyús Ádámmal), BUILD International, (1971) 58–63.
- *Panel Joints Prefabrication and Placing Tolerances* (Ottmár Bélával és Pattantyús Ádámmal), Building Science 4, (1972) 229–231.
- *Up-To-Date Design Methods of Outer Walls* (Ottmár Bélával, Pattantyús Ádámmal és Széll Máriával), Periodica Polytechnica, Vol. 19. No. 12, (1975)
- *Evaluation of Design* (Ottmár Bélával és Széll Máriával) Periodica Polytechnica, Vol. 19. No. 2, (1975)
- *Az információ divergencia egy alkalmazása az épületszerkezetek értékelésénél* (Ottmár Bélával), BME Bicentenárium Építésmérnöki Karon rendezett Tudományos Ülésszak előadásai, (1984) 65–68.
- *Többszempontú értékelési módszerek alkalmazása az építőipari beruházások előkészítésében* (Mályusz Leventével), Építéstechnológia - építési menedzsment '97, III (Döntéshozatal) témakör, (1997) 369–375.
- *Rácsos tartók állapotvizsgálata* (Kas Péterrel és Mályusz Leventével), Új utak a magyar operációkutatásban (szerk.: Komlósi Sándor, Szántai Tamás), Dialóg Campus Kiadó, Budapest-Pécs, (1999) 305–323.

Lineáris programozással és a pivot technikával kapcsolatos munkák

- *Magyar Módszer típusú algoritmusok Lineáris Programozási feladatok megoldására* (Terlaky Tamással), Alkalmazott Matematikai Lapok 12, (1985) 1–14.
- *Irányított matroidok megengedettségi feladatának egy új megközelítése* (Terlaky Tamással), Alkalmazott Matematikai Lapok 12, (1986) 279–282.
- *A Bland szabály a primál és a duál szimplex módszer esetén* (Terlaky Tamással), Alkalmazott Matematikai Lapok 13, (1987) 1–7.
- *Az ellipszoid módszerről* (Terlaky Tamással), SZIGMA XX, (1987) 196–208.
- *Remarks on the Feasibility Problem of Oriented Matroids* (Terlaky Tamással), Annales Universitatis Scientiarum Budapestiensis de Rolando Eötvös Nominata, Sectio Computatoria VII., (1987) 155–157.
- *A Criss-Cross módszer és az irányított matroidok komplementaritási feladata* (Terlaky Tamással), Alkalmazott Matematikai Lapok 14, (1989) 365–375.
- *A pivot technika szerepe a lineáris algebra néhány alapvető tételének bizonyításában* (Terlaky Tamással), Alkalmazott Matematikai Lapok 14, (1989) 425–448.

- *El Papel de la Technica del Pivote en la Demonstracion de Algunos Teoremas Fundamentales del Algebra Lineal* (Terlaky Tamással), *Revista de Investigacion Operational X.*, (1989) 141–167.
- *Some Generalizations of the Criss-Cross Method for Linear Complementary Problem of Oriented Matroids* (Terlaky Tamással), *Combinatorica* **9**, (1989) 189–198.
- *Variants of the Hungarian Method for Solving Linear Programming Problems* (Terlaky Tamással), *Optimization* **20**, (1989) 79–91.
- *The Role of Pivoting in Proving Some Fundamental Theorems of Linear Algebra* (Terlaky Tamással), *Linear Algebra and its Applications* **151**, (1991) 97–118.
- *On the Ellipsoid Method* (Terlaky Tamással), *Radovi Matematicki* **8**, (1992) 269–280.
- *Young Programming, an Analytical Approximation of Linear Programming* (Kas Péterrel, Mályusz Leventével és Gökhan Izbirakkal), *Research Report, Eastern Mediterranean University, EMU-AS-11* (1997)
- *Az exponenciális barrier programozás, mint a lineáris programozás analitikus megközelítése* (Mályusz Leventével), *Alkalmazott Matematikai Lapok* **19**, (1999) 199–216.
- *Approximation of Linear Programs by Bergman's DF Projection* (Kas Péterrel, Mályusz Leventével és Gökhan Izbirakkal), *European Journal of Operational Research* **126**, (2000) 69–79.
- *On the Dual of Linear Inverse Problems* (Kas Péterrel), *European Journal of Operational Research* **91**, (1996) 634–639.
- *A lineáris programozás analitikus megközelítése Young programozással* (Kas Péterrel és Mályusz Leventével), *Új utak a magyar operációkutatásban*, (szerk.: Komlósi Sándor, Szántai Tamás), *Dialóg Campus Kiadó, Budapest-Pécs*, (1999) 144–167.
- *Equilibrium Conditions of Trusses* (Kas Péterrel és Mályusz Leventével), *Publ. Univ. of Miskolc, Series D, Natural Sciences*, Vol. **39**, *Mathematics*, (1999) 47–56.

Csatorna kapacitási feladattal kapcsolatos munkák

- *A Geometric Programming Approach to the Channel Capacity Problem* (Mayer Jánossal és Terlaky Tamással), *Report, Department Operations Research, Eötvös University of Sciences No. 2*, (1991) 21 o.
- *A Geometric Programming Approach to the Channel Capacity Problem* (Mayer Jánossal és Terlaky Tamással), *Engineering Opt.* **19**, (1992) 115–130.

Hálózatokkal kapcsolatos munkák

- *Legrövidebb út meghatározása időtől függő élhosszal bíró hálózatban*, MTA Számítástechnikai Központ, Közlemények **3** (1967)
- *Determination of Shortest Path in a Network with Time-Dependent Edge-Lengths*, Math. Operationsforsch. u. Statist. **3**, (1972) 255–257.
- *An Algorithm for Solving the CPMtime-cost Trade-Off Problem* (Hajdú Miklóssal), Periodica Polytechnica Vol. **37**, No. **3**, (1993) 231–247.
- *Egy algoritmus a költségtervezési feladat megoldására tevékenység-élű term. hálón (CPM-cost feladat)* (Hajdú Miklóssal), Alkalmazott Matematikai Lapok **17**, (1993) 211–223.
- *Hálós tervezési technikák az építések tervezésében és irányításában* (Hajdú Miklóssal), Egyetemi jegyzet, Műegyetemi Kiadó, (1994) 183 o.
- *Optimization Techniques for Planning Highway Pavement Improvements* (Bakó Andrással, Gáspár Lászlóval és Szántai Tamással), Annals of Operations Research **58**, (1995) 55–66.

Statisztikai paraméterbecslési módszerekkel kapcsolatos munkák

- *Megjegyzés egy több dimenziós Cauchy eloszlásról* (Kas Péterrel), Alkalmazott Matematikai Lapok **13**, (1987) 145–161.
- *A több dimenziós Dirichlet eloszlás momentum és likelihood illesztésének numerikus megoldásáról* (Grubert Lászlóval), Alkalmazott Matematikai Lapok **15**, (1990) 197–211.

Young programozással kapcsolatos munkák

- *Convex Programs Based on the Young Inequality and its Relation to Linear Programming* (Kas Péterrel és Mályusz Leventével), Central European Journal of Operations Research **7**, (2000) 291–304.
- *Young Programming* (Kas Péterrel, Mályusz Leventével és Gökhan Izbirakkal), Encyclopedia of Optimization (szerk.: C. A. Floudos, R. M. Pardalos) (2001)

Előrebecslési és értékelő modellekkel kapcsolatos munkák

- *Az input-output tábla előrebecsléséről*, MTA Számítástechnikai és Automatizálási Intézet, Közlemények **10**, (1973)
- *A Theoretical Prediction of the Input-Output Tables*, I.F.I.P. TC7 Optimization Conference, 5th Conference on Optimization Techniques, (1973) 471–483.
- *A Theoretical Prediction of the Input-Output Table*, Lecture Notes in Computer Science **3**, (1973) 484–492.

- *Estimation of the Parameters in the Gravity Model for Trip Distribution, A New Method and Solution Algorithm* (Kádas Sándorral), *Regional Science and Urban Economics* **6**, (1976) 439–457.
- *Equivalence of Certain Types of Gravity and Entropy Maximization Models for the Trip Distribution Problem, An Application of Geometric Programming* (Kádas Sándorral), *Contributions to the Theory of Optimization* **2**, (1983) 38–70.
- *Gravitációs- és entrópia maximalizálási modellek* (Kádas Sándorral), *Adat - Modell - Elemzés* (szerk.: Kovács Erzsébet), *Aula Kiadó*, (2001) 37–52.
- *Linearly Constrained Estimation by Mathematical Programming* (Mayer Jánossal és Terlaky Tamással), *European Journal of Operational Research*, (1989) 254–267.
- *Hölder eltérés és alkalmazása a többtényezős értékelés feladataiban*, *PRO-DIFORM Tanulmány* (1992)

Egyéb munkák

- *A keverési feladat matematikai modelljeiről* (Mayer Jánossal és Terlaky Tamással), *Alkalmazott Matematikai Lapok* **14**, (1989) 99–117.
- *A New Convergent Algorithm for the Continuous Modular Design Problem* (Mayer Jánossal és Terlaky Tamással), *The Arabian Journal for Science and Engineering* **15**, (1990) 687–694.
- *A Proof of the Generalized Hadamard Inequity via Information Theory* (Kas Péterrel), *Annales Universitatis Scientiarum Budapestiensis de Rolando Eötvös Nominata, Sectio Computatorica XIII*, (1992) 21–24.
- *Some Generalizations of the Criss-Cross Method for Quadratic Programming* (Terlaky Tamással), *Optimization* **24**, (1992) 127–139.
- *On the Duality of the Mixed Entropy Programming* (Kas Péterrel), *Optimization* **27**, (1993) 253–258.
- *A New Algorithm for the Continuous Modular Design Problem* (Terlaky Tamással és Mayer Jánossal), *Proceedings of the SIGAL Workshop on Algorithms* **10**, Japan, Association of Information Science (2001) 1–8.
- *Sztochasztikus jelenségek* (Nagy Tamással), *Operációkutatás*, Vol. 1, *Budapesti Közgazdaságtudományi és Államigazgatási Egyetem, Aula Kiadó* (2002)

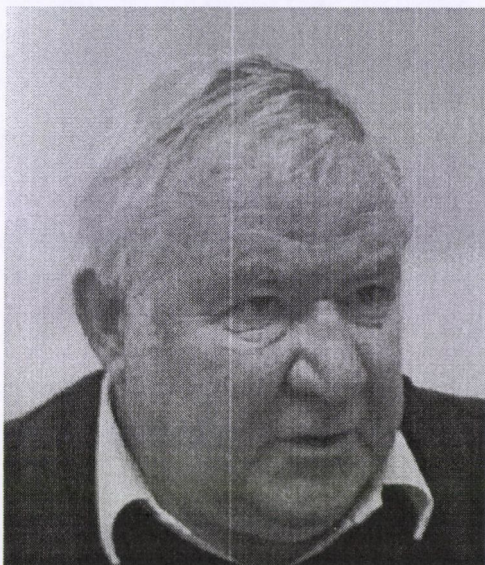
Kutatási pályázati munkák

- *EGPO 59/86 kutatási pályázat*, 1986–1989.
 - A kvadratikus programozás két alapvető algoritmus, 1987.
 - A kvadratikus programozás dualitási problémaköre és a megoldására szolgáló Lemke algoritmus és komputer kódja, 1988.

- A kvadratikus programozási feladat megoldására szolgáló Van de Panne algoritmus és komputer kódja, 1988.
- A sorozatgyártási és a modulervezési feladat, 1989.
- *OTKA kutatási pályázat*, 1991–1992.
 - Egy primál-duál algoritmus az entrópia programozási feladat megoldására, 1991.
 - A nemnegatív vektorok eltérésének mérésére szolgáló eltérésfüggvények 3 típusa, 1992.
 - A nemnegatív vektorok eltérésének mérésére szolgáló eltérésfüggvények vizsgálata és alkalmazása a döntési modellekben, 1993.
- *FKFP 0231 kutatási pályázat*, 1999–2002.
 - Sztochasztikus jelenségek, 2000.

Összeállította: Nagy Tamás (Miskolci Egyetem)

HARNOS ZSOLT (1941–2009)



Harnos Zsolt szakmai életrajza

Harnos Zsolt 1941. március 27-én született Budapesten. Az ELTE TTK alkalmazott matematikus szakán 1966-ban szerzett diplomát.

1966 és 1987 között az MTA, az OMFB és az OT különböző intézeteiben dolgozott. 1987-től a Kertészeti és Élelmiszeripari Egyetem (2000-től Szent István Egyetem, majd 2003-tól Budapesti Corvinus Egyetem) Matematika és Informatika tanszék tanszékvezető egyetemi tanára. 2000 és 2006 között a Szent István Egyetem, illetve a Budapesti Corvinus Egyetem tudományos rektorhelyettese volt.

1978-ban nyerte el a matematikai tudományok kandidátusa fokozatot (Szubderiváltak és érintőkúpok vizsgálata és alkalmazásai optimalizációs problémákra). 1985-ben kapta meg a tudományok doktora fokozatot (Az agroökológiai adottságok rendszerének matematikai modellezése).

1995-ben a Magyar Tudományos Akadémia levező tagjává választották. Székfoglaló előadását Informatika szerepe az agrártudományokban címmel tartotta meg.

2001-ben a Magyar Tudományos Akadémia rendes taggá választotta. Székfoglaló előadásának címe: Mezőgazdasági rendszerek modellezésének néhány problémája.

1997-ben akkreditálták az Élelmiszergazdaság döntéstámogató rendszerei című doktori programot. Ez 2001-ben Doktori Iskolává alakult, melynek vezetője volt.

Az elmúlt években tíz doktorandusz témavezetője volt, akik közül többen már megszerezték a PhD fokozatot. Tudományos rektorhelyettesként felügyelte a SzIE, illetve 2003-tól a Budapesti Corvinus Egyetem doktorképzését. A Budapesti Corvinus Egyetemen a Tudományos Tanács és az Élettudományi Doktori Tanács elnöke volt.

Számos szakmai és társadalmi funkciót töltött be a KÉE, SzIE, illetve BCE bizottsági rendszerében.

Harnos Zsolt **kutatási területei** az alkalmazott informatika fejlődésével, általánossá válásával párhuzamosan jelentős változásokon mentek keresztül. A 70-es évek közepe óta foglalkozott a közgazdaság, agrárgazdaság, ill. a környezetvédelem, környezetgazdálkodás matematikai, informatikai kérdéseivel.

Matematikai módszertani, alkalmazott informatikai kutatásai nagyrészt jelentős hazai és nemzetközi kutatási projektekhez kapcsolódtak. Ezek közül kiemelkedők

- a magyar mezőgazdaság agroökológiai potenciáljának a felmérése,
- a biomassza hasznosítási lehetőségeinek a feltárása,
- az aszályok elemzése,
- az AGRO'21,
- az AGRO-QUALITÁS programok,
- az agrotechnika termőképességre gyakorolt hatásainak az elemzése (IIASA – nemzetközi projekt),
- a klímaváltozás és annak várható hatása a mezőgazdaságra (CLAIRE, CLIVARA).

A 90-es évek óta kutatási tevékenysége elsősorban az időjárási változékonyság, a klímaváltozás előrejelzett alakulásának elemzésére, annak a mezőgazdaságra gyakorolt hatásának feltárására irányult.

A kutatások két egymásra épült nemzetközi kutatási projekt keretében folytak: CLAIRE és CLIVARA projekt, az Environment Change Unit, Oxford University koordinálásában 11 ország, s azon belül 16 kutatóintézet közreműködésével, a Közép-Európai régiót a Harnos Zsolt által vezetett csoport képviselte.

Részt vett a VAHAVA (VÁLtozás-HATás-VÁlaszadás) elnevezésű projektben 2003 és 2006 között, amelynek folytatását „Felkészülés a klímaváltozásra: Környezet-Kockázat-Társadalom (KLIMAKKT)” címmel vezette.

2007-ben vezetésével alakult az Alkalmazkodás a klímaváltozáshoz elnevezésű akadémiai kutatócsoport és az MTA-n a Klímavédelmi Kutatások Koordinációs Iroda.

Vezetésével vett részt egy kutatócsoport az ADAM – ADaptation And Mitigation Strategies: supporting European climate policy elnevezésű Európai Projektben.

A 80-as évek vége óta elnöke a Biometriai és Biomatematikai Bizottságnak és az Operációkutatási Bizottságnak (1999-ig). Tagja volt a Környezettudományi Elnöki Bizottságnak, a Könyvtári Bizottságnak, az Aszály Bizottságnak (alelnök). 1994–1998 elnöke a Nemzetközi Biometriai Társaság (IBS) Magyar Régiójának, 2002-ben újra elnöki megbízást kapott. 1996–1998 elnöke volt a Magyar Operációkutatási Társaságnak. Elnökségi tagja volt a HUNINET és HUNGARNET egyesületnek. 2003-tól 2006-ig tagja volt az OTKA Bizottságnak.

Az elmúlt évtizedben folyamatosan részt vett több nemzetközi szervezet munkájában. Az IFORS (International Federation on Operation Research Societies) magyar képviselőjeként, valamint az IBS (International Biometrical Societies) council tagjaként.

A Nemzetközi Biometriai Bizottság 1990-ben rendezte Magyarországon XV.-ik konferenciáját, az EURO (az IFORS európai szervezete) 2000-ben tartotta Magyarországon nemzetközi konferenciáját, amelyek szervezésében Harnos Zsolt jelentős szerepet vállalt. 2003-ban az EFITA magyarországi konferencia programbizottságának elnöke. A fentiekén kívül több hazai és nemzetközi konferencia szervezésében vett részt.

1995 és 2000 között tagja volt a JABES (Journal of Agricultural, Biological and Environmental Statistics) és 1988 óta tagja volt az Alkalmazott matematikai lapok című folyóiratok szerkesztőbizottságának. Tagja volt a Gazdálkodás és az AGRO21 füzetek szerkesztő bizottságának.

1999 és 2002 között Széchenyi Professzori Ösztöndíjban részesült. 2002-ben kutató és oktató munkája elismeréseként megkapta a Magyar Köztársasági Érdemrend Tisztikeresztje kitüntetést.

2006-ban a tudományos utánpótlás nevelésben, az élettudományi doktori iskolák létrehozásában és szervezésében kifejtett iskolateremtő tevékenysége elismeréseként Széchenyi-díjat kapott.

Harnos Zsolt publikációi

Könyvek, könyvrészek

- [K1] HARNOS ZS.: *ICL System 4 ALGOL gépi reprezentáció*, OTSZK, (1971) 90 old.
- [K2] HARNOS ZS.: *Differenciálszámítás és általánosításai*, Magyar Rendszeranalízis Bizottság Kiadványa, (1979) 75 old.
- [K3] HARNOS ZS.: *Az ökológiai feltételek matematikai modellezése*, Az MTA és a Bolgár Tudományos Akadémia első együttes szemináriuma az élelmiszeripari komplexum témakörben, Budapest, (1981) (orosz nyelven)
- [K4] LÁNG I., CSETE L., HARNOS ZS.: *A magyar mezőgazdaság agroökológiai potenciálja az ezredfordulón*, Mezőgazdasági Kiadó, Budapest, (1983) 265 old.

- [K5] CSÁKI, CS., HARNOS, ZS., LÁNG, I.: *Agricultural Development and Ecological Potential: The Case of Hungary*, Kieler Wissenschaftsverlag Vauk, (1984) pp. 130
- [K6] LÁNG I., HARNOS ZS., CSETE L., KRALOVÁNSZKY U. P., TÓKÉS O.: *A biomassza hasznosításának lehetőségei*, Mezőgazdasági Kiadó, Budapest, (1985) 249 old.
- [K7] LÁNG, I., CSÁKI, CS., CSETE, L., HARNOS, ZS., CSUMACSENKO, B. A., VLASZOV, E. P., LADUNGA, I.: *A KGST tagországok agráripari komplexumainak rendszerelemzése és modellezése: A magyar esettanulmány*, Moszkva, MNIIPU, (1986) (orosz nyelven) 170 old.
- [K8] HARNOS, ZS.: *Agricultural models (Ch. 11)*, in Leon Braat and Wal van Lierop (Eds.) *Economic - ecological modeling*, North-Holland, (1987) 100–127.
- [K9] LÁNG I., HARNOS ZS.: *Aszály és kockázat*, Rendszerkutatói tanulmányok 4. kötet Akadémiai Kiadó, Budapest, 1987
- [K10] CSÁKI, CS., HARNOS, ZS., RAJKAI, K. AND VÁLYI, I.: *Hungarian agriculture: Development potential and environment (Ch. 9)*, in J. K. Parikh (Ed.) *Sustainable Development in Agriculture*, Martinus Nijhoff Publishers, (1988) 253–297.
- [K11] HARNOS ZS.: *A módszertani kutatások eredményei (II. fejezet)*, in Láng I., Csete L. (szerk.) *Az alkalmazkodó mezőgazdaság*, AGRICOLA, Budapest, (1992) 12–44.
- [K12] HARNOS ZS.: *Időjárás és időjárás-termés összefüggéseinek idősoros elemzése*, in Baráth Csabáné, Györffy B., Harnos Zs. (szerk.) (I. fejezet) *Aszály* (1983), AKAPRINT, Budapest, (1993) 9–46.
- [K13] HARNOS ZS. (TÁRSSZERZŐ): *Biometriai módszerek és alkalmazásai MINITAB programcsomaggal*, AKAPRINT, Budapest, (1993) 240 old.
- [K14] HARNOS ZS.: *A klímaváltozás egy előrejelzése és a magyarországi tendenciák Éghajlat, időjárás, aszály*, AKAPRINT, Budapest, (1994) 23–58.
- [K15] KARACOSTAS, T. S., HARNOS, ZS.: *Agriculturally Significant Climatic Extremes in Climatic Change and Agriculture in Europe: Assessment of Impacts and Adaptation* Environmental Change Unit, University of Oxford, (1994) 78–105.
- [K16] HARNOS, ZS.: *Crop-specific climatic extremes for wheat and maize in Hungary in Climatic Change and Agriculture in Europe Assessment of Impacts and Adaptations*, Eds: P. A. Harrison, R. E. Butterfield and T. E. Dawning ECU, Univ. of Oxford (1995) 105–126.
- [K17] HARNOS, ZS. AT AL: *Risk modelling: SALINEX in Agroclimatic Change and European Soil Suitability*, Eds. P. J. Loveland and M. D. A. Ronnsevell Craunfield University, Silsoe, (1996) 95–122.

- [K18] MOLNÁR, S., HARNOS, ZS. AT AL: *Hungarian Climate Change Country Study*, Systemexpert Cons. Ltd., Budapest, (1996) pp. 121
- [K19] HARNOS, Zs.: *Modelling crop response in Hungary* in Climate Change, Climate Variability and Agriculture in Europe, Eds. P. A. Harrison, R. E. Butterfield and T. E. Downing, Oxford, (1996), Environmental Change Unit, University of Oxford, 179–190.
- [K20] HARNOS, Zs.: *Élelmiszeripar és informatika* in Magyarország az ezredfordulón, A magyar agrárgazdaság jelene és kilátásai Szerk. Glatz Ferenc, MTA, Budapest, (1997) 174–176.
- [K21] HARNOS, ZS. AT AL.: *Hungarian climate change action plan*. Published by Systemexpert Consulting Ltd., Budapest, (1998) pp. 93
- [K22] HARNOS, ZS., BUSSAY, A., HARNOS, N.: *Modelling climate change impacts on wheat and potato in Hungary (Ch. 19)* in Climate change, climatic variability and agriculture in Europe, Eds. P. A. Harrison, R. E. Butterfield and T. E. Downing, Environmental Change Unit, University of Oxford (1999) pp. 16, 249–260.
- [K23] HARNOS, Zs.: *Informatika szerepe a minőségbiztosításban*, Az agrárgazdaság minőségi dimenziói, Láng I. és Csete L. (szerk.) MTA Stratégiai Kutatások, (1999) 83–97.
- [K24] HARNOS, ZS., KÁRPÁTI, L., SZENTELEKI, K.: *Informatika az agrárgazdaságban*, Az agrárgazdaság közgazdasági, vállalkozási és szabályozási ismeretei II., Szőlő-bor ágazat szakirányú kérdései, NYITSAK (1999) 67–114.
- [K25] HARNOS ZS.: *Az informatika szerepe a minőségbiztosításban*, Glatz Ferenc (szerk.), Minőség és agrárstratégia, MTA Budapest (1999)
- [K26] HARNOS, ZS., BUSSAY, A., HARNOS, N.: *Modelling climate change impacts on wheat and potato in Hungary (Ch. 19)* in Climate change, climatic variability and Agriculture in Europe, Eds. T. E. Downing, P. A. Harrison, R. E. Butterfield and K. G. Lonsdale, Environmental Change Unit, University of Oxford, (2000) 349–367.
- [K27] HARNOS, ZS. AT AL *Review and comparison of scaling-up methods* in Climate change, climatic variability and Agriculture in Europe, Eds. T. E. Downing, P. A. Harrison, R. E. Butterfield and K. G. Lonsdale, Environmental Change Unit, University of Oxford (2000) 393–415.
- [K28] HARNOS ZS.: *Az agroökológiai potenciál felmérésének értékelése*, Környezetügy 2004, Tanulmányok Láng István tiszteletére, Országos Környezetvédelmi Tanács Friedrich Ebert Alapítvány, Budapest, (2004) 121–134.
- [K29] HARNOS ZS., HUFNAGEL L.: *Adatértékelési módszerek és a precíziós gazdálkodás döntéstámogató rendszerei*, Németh, T., Neményi M., Harnos Zs. (szerk.): A precíziós mezőgazdaság módszertana, JATEPress - MTA TAKI, Szeged (2007) 159–204.

- [K30] HARNOS Zs.: *A kockázatelemzés módszertani kutatásainak összefoglaló áttekintése*, Harnos, Zs., Csete, L. (szerk): *Klimaváltozás: környezet – kockázat – társadalom*, Szaktudás Kiadó Ház, Budapest, (2008) 267–275.
- [K31] HARNOS Zs. – GAÁL M. – HUFNAGEL L.: *Klimaváltozásról mindenkinek*, Budapesti Corvinus Egyetem MIT, Budapest, (2008) 200 old.

Szerkesztett könyvek

- [SZ1] HARNOS, Zs.: *The XVth International Biometric Conference Proceedings, Contributed Papers*, Budapest, (1990)
- [SZ2] HARNOS, Zs.: *The XVth International Biometric Conference Proceedings, Invited Papers*, Budapest, (1990)
- [SZ3] HARNOS Zs.: *Az alkalmazkodó mezőgazdaság rendszere: Módszertani kutatások*, AKAPRINT, Budapest, (1991)
- [SZ4] HARNOS Zs.: *Biometriai módszerek és alkalmazásai MINITAB programcsomaggal*, AKAPRINT, Budapest, (1993)
- [SZ5] Baráth Csabáné, Györffy B., Harnos Zs.: *Aszály 1983*, AKAPRINT, Budapest, (1993)
- [SZ6] CSELŐTEI LÁSZLÓ, HARNOS ZSOLT: *Éghajlat, időjárás, aszály - I. Az időjárás változékonysága és hidrológiai vonatkozásai*, AKAPRINT, Budapest, (1994)
- [SZ7] HARNOS, Zs.: *EN+IN, International Conference on Environment and Informatics*, Conference Proceedings Budapest, (1995)
- [SZ8] CSELŐTEI L., HARNOS Zs.: *Éghajlat, időjárás, aszály - II. Az aszály enyhítésének lehetőségei* AKAPRINT, Budapest, (1996)
- [SZ9] HARNOS Zs.: *AGRÁRINFORMATIKA '99*, Informatikai kutatások, fejlesztések és alkalmazások az agrárgazdaságban, Konferencia kiadvány, DATE 1999.
- [SZ10] HARNOS Zs., KOVÁCS L. M.: *Az agrárágazat közigazgatási, vállalkozási és szabályozási ismeretei I-II.*, Szőlő és bor ágazat szakirányú kérdései, NYITSZAK (1999)
- [SZ11] Zs. HARNOS, M. HERDON, T. B. WIWCZAROSKI: *EFITA 2003 - Information technology for a better agri-food sector, environment and rural living*, Proceedings University of Debrecen, Hungarian Association of Agricultural Informatics, Debrecen-Budapest, (2003) pp. 909
- [SZ12] HARNOS Zs., HREDON M.: *Summer University on Information Technology in Agriculture and Rural Development*, University of Debrecen, Hungarian Association of Agricultural Informatics, Debrecen, (2006) pp. 184
- [SZ13] NÉMETH, T., NEMÉNYI M., HARNOS Zs.: *A precíziós mezőgazdaság módszertana*, JATEPress - MTA TAKI, Szeged (2007)

- [SZ14] HARNOS ZS. – CSETE L.: *Klímaváltozás: környezet – kockázat – társadalom*, Kutatási eredmények, Szaktudás Kiadó Ház, Budapest, (2008) 380 old.
- [SZ15] HARNOS ZS. – GAÁL M. – HUFNAGEL L.: *Klímaváltozásról mindenkinek*, Budapesti Corvinus Egyetem MIT, Budapest, (2008) 200. o.

Tankönyv, egyetemi jegyzet

- [J1] HARNOS ZS.: *Operációkutatási módszerek és alkalmazásai*, Kertészeti és Élelmiszeripari Egyetem, Budapest, (1990) 204 old.
- [J2] HARNOS ZS.: *Matematika II.*, Kertészeti és Élelmiszeripari Egyetem, Budapest, (1998) 140 old.
- [J3] HARNOS, ZS., KÁRPÁTI, L., SZENTELEKI, K.: *Informatika az agrárágazatban*, Az agrárágazat közigazgatási, vállalati és szabályozási ismeretei II., (Szerk: Harnos Zsolt és Kovács László Miklós), NYITSAK Budapest, (1999) 47 old.
- [J4] PROF. HARNOS ZSOLT – LADÁNYI MÁRTA: *Biometria agrártudományi alkalmazásokkal*, Aula Kiadó, (2005) 338 old.

Tanulmányok

- [DP1] HARNOS ZS.: *A differenciálszámítás és néhány általánosításának geometriai vizsgálata Banach terekben, s azok alkalmazása az optimalizációban* (doktori disszertáció) Budapest (1976)
- [DP2] HARNOS ZS.: *Szubderiváltak és érintőkúpok vizsgálata és alkalmazásai optimalizációs problémákra* (kandidátusi disszertáció) Budapest (1978)
- [DP3] LÁNG I., CSÁKI CS., HARNOS ZS., LADUNGA I.: *Az agroökológiai potenciál felmérése Magyarországon és prognózisa a 2000. évig*, Moszkva (1984) (orosz nyelven)
- [DP4] LÁNG I., CSETE L., HARNOS ZS.: *A biomassa fokozottabb hasznosítása*, Nemzetközi Mezőgazdasági Szemle, XXVIII. évf. (1984)
- [DP5] HARNOS ZS.: *A mezőgazdasági termelés környezeti hatásainak modellezése* (1984) „A környezetvédelem prognosztizálásának és tervezésének kérdései”, K-5/VII. főtéma
- [DP6] CSÁKI CS., FORGÁCS CS., HARNOS ZS., MÓDOS GY., RAJKAI K., SEBESTYÉN M.: *Matematikai modell a talaj-növény kapcsolatok, valamint a vállalati termőföld-hasznosítási rendszer vizsgálatára* (két kötet), MKKE Agrárgazdasági Tanszék (1984)
- [DP7] HARNOS ZS.: *A biológiai eredetű anyagok (biomassa) hasznosításának távlati lehetőségei*. A Komplex Bizottság jelentése II. kötet: A modellrendszer, MTA (1984)

- [DP8] LÁNG, I., HARNOS, ZS., MADAS, A.: *Hungarian Biomass Resources*, Great Britain - East Europe Centre Biomass Round Table, Budapest (1984)
- [DP9] HARNOS ZS.: *Az agroökológiai adottságok rendszerének matematikai modellezése* (MTA doktori értekezés), Budapest (1985)
- [DP10] LÁNG I., HARNOS ZS.: *Környezetkímélő mezőgazdasági rendszerek*. A Balaton Csoport, s a Magyar Agrártudományi Egyesület budapesti tudományos értekezletén elhangzott előadás készült cikk, Búvár, XV. évf., 1 (1985)
- [DP11] HARNOS, ZS.: *Global Aspects of Biomass Production and Utilization*, Pozsony (1987. július)
- [DP12] HARNOS ZS. ET AL.: *Az élelmiszergazdaság és az infrastruktúra fontosabb elemeinek kapcsolata: Az élelmiszergazdaság információs rendszere*, OMFB 9-8901/7-I tanulmány (1990)
- [DP13] HARNOS ZS., SZVIREZSEV, J. M., RACSKÓ P., SZEIDL L., SZEMENOV, M. A.: *A gabonatermesztés rendszerelemzése és termésbecslés*, Budapest (1991. augusztus)
- [DP14] HARNOS ZS.: *Az ökológiai alapú tájtermelés tervezésének módszertani eszközei*, Földművelésügyi Minisztérium (1992)
- [DP15] CZELNAI R., HARNOS ZS., SZÁSZ G.: *Éghajlatváltozás és aszály: Kutatási eredmények, irányzatok és magyarországi feladatok*, MTA Aszálybizottság (1994)
- [DP16] HARNOS ZS., FERENCZY A.: *A szolnoki régió időjárásának jellemzése*, Budapest (1994)
- [DP17] HARNOS ZS.: *A klímaváltozás egy előrejelzése és a magyarországi tendenciák*, Budapest (1994)
- [DP18] HARNOS ZS.: *A kutatás, oktatás és szaktanácsadás szerepe a fenntartható agrárgazdaságban*, A fenntartható fejlődés időszerei kérdései a mezőgazdaságban. XXXVII. Georgikon Napok Keszthely (1995), Konferencia kiadvány, 279–284.
- [DP19] HARNOS ZS.: *BORINFO: Információs rendszerek a szőlő-borgazdaságban*, IX. Országos Szőlész-Borász továbbképző tanfolyam, Balatonfüred, (1997) 53–56.
- [DP20] HARNOS ZS., SZENTELEKI K.: *Szőlő-bor ágazat információs rendszere*, XXI. Országos Szőlész-Borász Továbbképző Tanfolyam, Balatonfüred, (1999) 64–68.
- [DP21] HARNOS ZS.: *Mezőgazdasági Biometria és Biomatematikai Bizottság*. Az Osztályhoz tartozó tudományos bizottságok kialakulása, célja, küldetése, MTA Agrártudományok Osztálya, Budapest, (1999 évkönyv) 233–239.

- [DP22] HARNOS Zs.: *Az Osztály akadémikusainak pályatükre*, MTA Agrártudományok Osztálya, Budapest, (1999) 637–646.
- [DP23] HARNOS Zs.: *Informatika szaktanácsadóknak*, Szaktanácsadási kiadvány (társszerző) FVM Budapest (2000)
- [DP24] HARNOS Zs.: *Az agrárfelsőoktatás és az agrárkutatások kapcsolatai*, MTA Agrártudományok osztályának 2000. évi tájékoztatója, Az MTA közgyűléséhez kapcsolódó nyilvános ülés, Korszakváltás és kihívások a magyar agrártudományokban, Budapest (2001) 121 old.
- [DP25] HARNOS Zs.: *Informatika a precíziós mezőgazdaságban*, EU konform mezőgazdaság és élelmiszerbiztonság, Debreceni Egyetem, ATC KITE Rt. MTA-DE Földművelési Kutatócsoport, Debrecen, (2002) 78–83.
- [DP26] HARNOS Zs.: *Mezőgazdasági rendszerek modellezésének néhány problémája*, Székfoglaló előadás (tömörített változat), MTA Agrártudományok Osztályának 2002. évi tájékoztatója, Budapest, (2003) 258–263.
- [DP27] Zs. HARNOS: *Risk handling in a decision support system*, Information technology for a better agri-food sector, environment and rural living, EFITA 2003 Conference, Debrecen-Budapest, (2003) 33–42.
- [DP28] HARNOS Zs.: *Az agroökológiai potenciál felmérésének értékelése*, Környezetügy 2004, Tanulmányok Láng István tiszteletére, Országos Környezetvédelmi Tanács Friedrich Ebert Alapítvány, Budapest, (2004) 121–134.

Könyvismertetés

- [KI1] BARÁTH CSABÁNÉ DR., ITTZÉS ANDRÁS, DR. UGRÓSDY GYÖRGY: *Biometria módszertani alapok a MINITAB programcsomag alkalmazásával*, Növénytermesztés, Tom. 46. No. 3., Budapest, (1997) 231–322.

Tudományos közlemények

- [T1] HARNOS, Zs.: *On Closedness of the Image of Convex, Closed Sets*, Annales Sci. Univ. Sec. Math. Tom. XIII., (1970)
- [T2] HARNOS Zs.: *Ártípuszámítások során felvetődő egyenletrendszerek numerikus megoldásáról*, Információ - Elektronika (1972)
- [T3] HARNOS, Zs.: *On Differentiability and Subdifferentiability of Series of Convex Functions*, ELTE TTK Szemelvények 1977/5
- [T4] HARNOS, Zs.: *On Differentiability of K-convex Functions and its Applications System*, Analysis and Related Topics, Mathematica No. 15., (1977)
- [T5] HARNOS, Zs.: *On the Geometric Characterization of Differentiation in Normed Vector Spaces*, Annales Sci. Univ. Sec. Math. Tom XXI., (1978)
- [T6] HARNOS, Zs.: *Divisors of Polynomials with Positive Coefficients*, Periodica Mathematica Hungarica, Vol. 11, 2, (1980)

- [T7] LÁNG I., HARNOS ZS.: *Tájékoztató „A magyar mezőgazdaság agroökológiai potenciáljának felmér” című munka néhány módszertani kérdéséről és fontosabb eseményéről*, Prognosztika, 3–4 szám, (1981)
- [T8] CSETE, L., HARNOS, ZS., LÁNG, I.: *Measuring the agroecological potential in Hungary and its expected trends up to 2000*, AKI Bulletin, No. 49., (1981)
- [T9] HARNOS, ZS.: *Long Range Impacts and Consequences of Technological Development In Hungarian Agriculture*, in Food for all in a Sustainable World: The IIASA Food and Agriculture Program, (J.K. Parikh and F. Rabár, Editors) IIASA, Laxenburg, (1981)
- [T10] HARNOS, ZS.: *The Mathematical Model for the Determination of the Optimal Crop. Production Structures*, „Large-Scale Linear Programming” Vol. 2. (1981) (G. B. Dantzig, M. A. H. Dempster and M. I. Kallio, Editors)
- [T11] HARNOS ZS.: *Természeti tényezők a rendszerszemlélet tükrében*, Pécs, (1981), Baranya megye agroökológiai potenciálja
- [T12] LÁNG I., CSETE L., HARNOS ZS.: *Magyarország agroökológiai potenciáljának felmérése és várható alakulása 2000-ig*, Gazdálkodás, XXV. évf. 1. sz., (1981)
- [T13] HARNOS, ZS.: *The Survey of the Agroecological Potential of Hungary - A brief summary*, IIASA, Laxenburg, CP-82-21, (1982)
- [T14] HARNOS, ZS., TÓTH, F.: *The Protein Problem in the Complex Framework of Biomass Utilization*, System Analytical Approach to the Assessment of Non-Conventional Protein Production Technologies, Proceeding of Task Force Meeting, Sofia, (1982)
- [T15] HARNOS, ZS., GYÖRFFY, B.: *Estimation of Crop Yield Potentials as Affected by Agroecological Conditions*, Environmental Aspects in Global Modeling, G. Bruckmann Editor, IIASA, Laxenburg, CP-82-21, (1982)
- [T16] HARNOS, ZS.: *The Mathematical Model for the Determination of the Optimal Crop Production Structures as Affected by Agroecological Conditions*, Modeling Agricultural-Environmental Processes, IIASA, Laxenburg, CP-82-S5, (1982)
- [T17] CSÁKI, CS., HARNOS, ZS., VÁLYI, I.: *Methodology for the Investigation of Long Term Consequences of Development in Hungarian Agriculture*, An IIASA FAP Task 2 Case Study, IIASA, Laxenburg, WP-82-62, (1982)
- [T18] LÁNG I., CSETE L., HARNOS ZS., SCHULTZ M.: *A biomassza fokozottabb hasznosítása*, Gazdálkodás, XXVII. évf., 1983/7
- [T19] LÁNG I., HARNOS ZS.: *A biomassza program főbb célkitűzései és modell-rendszere*, Természeti erőforrások, (1983)
- [T20] LÁNG, I., HARNOS, ZS., CSETE, L.: *Possibilities of Biomass Production and Utilization in Hungary*, Agrokémia és Talajtán, Tom. XXXIV (1985)

- [T21] LÁNG, I., HARNOS, ZS.: *Economic and Social Constraints in Establishing Sustainable Agricultural Systems in Hungary*, Agrokémia és Talajtan, Tom. XXXIV (1985)
- [T22] HARNOS, ZS.: *Resources and Utilization of Lignocellulosic Biomass in Hungary*, Report on the 1st EC Workshop on „Agricultural Surplusses”, Frankfurt/Main, (1985)
- [T23] HARNOS, ZS.: *Resources and Utilization of the Starch Containing Plants in Hungary*, Report on the 1st EC Workshop on „Research and Development”, Frankfurt/Main, (1985)
- [T24] LÁNG, I., HARNOS, ZS., NAGY, Á., VÁLYI, I.: *The Biopotential of Hungary*, International Agrophysics, 1, (1985)
- [T25] HARNOS, ZS. – CSÁKI, CS.: *Methodology for the Investigation of Long Term Consequences of Technological Development in Hungarian Agriculture and the Management of Natural Resources*, Proceedings of Eighth Symposium of EAAE, F. Lechi Editor, Kieler Wissenschaftsverlag, Vauk, (1985)
- [T26] HARNOS, ZS.: *Perspectives for Developing the Land Component of the Biosphere Program*, IIASA, Laxenburg, CP-86-22, (1986)
- [T27] CSÁKI, CS., HARNOS, ZS., RAJKAI, K., VÁLYI, I.: *Development Potentials and Environment in Hungarian Agriculture*, Report on a Hungarian-IIASA Case Study, IIASA, (1986)
- [T28] HARNOS, ZS.: *A Land Evaluation Methodology and Its Possible Use in Long Term Planning*, Vth European Congress of Agricultural Economists (1987), Working Group B 3 Papers
- [T29] LÁNG, I., CSETE, L., HARNOS, ZS.: *The Enterprisal System of an Adjusting Agriculture in Hungary*, European Review of Agricultural Economics, Vol. 15., No. 2/3, Proceedings Fifth Congress of the E. A. A. E. Resource, Adjustment and European Agriculture, (1988) 225–239.
- [T30] VARGA-HASZONITS, Z., HARNOS, ZS.: *Effect of Climate Variability and Drought on Wheat and Maize Production*, in Identifying and Coping with Extreme Meteorological Events, (Eds. Antal and Glantz), Hungarian Meteorological Service, (1988)
- [T31] HARNOS ZS.: *Az időjárás változékonysága és hatása a növénytermesztésre*, Vízügyi Közlemények, LXX. évf., 3. sz., (1988) 397–408.
- [T32] HARNOS ZS.: *Az operációkutatás helye és szerepe a módszertani kutatások körében*, Szigma, (1988)
- [T33] HARNOS, ZS. *Methodology of Hierarchical Multiobjective Optimization and its Use in the Planning of Agricultural Development*, Lectures Notes in Economics and Mathematical Systems, Springer-Verlag, (1989) 127–146.

- [T34] LÁNG, I., CSETE, L., HARNOS, ZS.: *L'adaptation du systeme des enterprises dans l'agriculture hongroise*, Economie Rurale, 189 (1989)
- [T35] HARNOS ZS.: *Az alkalmazkodó mezőgazdaság egy rendszerelemzési megfogalmazása*, Gazdálkodás, XXXIII. évf. 1989. 2. sz.
- [T36] HARNOS ZS.: *Ökonómiai eszközök az ökológiai folyamatok szabályozásában*, Az Európai Agrárközgazdászok V. Kongresszusának tanulságai, (Balatonszéplak, 1987), Agroinform, (1989) 37–42.
- [T37] HARNOS ZS.: *Sztochasztikus optimalizáció alkalmazása a mezőgazdaságban*, XXXII. Georgikon Tudományos Napok, (1990)
- [T38] HARNOS ZS.: *Az alkalmazkodó mezőgazdaság rendszere: módszertani alapok*, Az alkalmazkodó mezőgazdaság rendszere c. könyvben, AKAPRINT, Budapest, (1991)
- [T39] HARNOS ZS., RACSKÓ P., SZEIDL L., VARGA-HASZONITS Z.: *Meteorológiai viszonyok elemzése*, Az alkalmazkodó mezőgazdaság rendszere c. könyvben, AKAPRINT, Budapest, (1991)
- [T40] FERENCZY A., HARNOS ZS., RACSKÓ P., SZEIDL L., VARGA-HASZONITS Z.: *Termés-időjárás kapcsolata*, Az alkalmazkodó mezőgazdaság rendszere c. könyvben, AKAPRINT, Budapest, (1991)
- [T41] BÁNKÖVI GY., HARNOS A., HARNOS ZS., LADÁNYI M., VELICZKY J.: *Adaptív előrejelzés*, Az alkalmazkodó mezőgazdaság rendszere c. könyvben, AKAPRINT, Budapest, (1991)
- [T42] HARNOS ZS., RACSKÓ P., SZEIDL L., SZEMENOV, M., SZVIREZSEV, J.: *Kockázatelemzés*, Az alkalmazkodó mezőgazdaság rendszere c. könyvben, AKAPRINT, Budapest, (1991)
- [T43] HARNOS ZS.: *A környezetkímélő mezőgazdaság tervezése*, Az alkalmazkodó mezőgazdaság rendszere c. könyvben, AKAPRINT, Budapest, (1991)
- [T44] GERENDAY Á., HARNOS ZS., RACSKÓ P., SZALAY E., SZENTELEKY K.: *Agroökológiai Integrált Információs Rendszer (AIIR)*, Az alkalmazkodó mezőgazdaság rendszere c. könyvben, AKAPRINT, Budapest, (1991)
- [T45] HARNOS ZS.: *Az aszály-felmérések áttekintése, következtetések, stratégiai megfontolások, megoldandó feladatok*; Az 1992. évi aszály értékelése, a Földművelésügyi Minisztériumban 1992. november 18-án tartott konferencia anyaga, EUROPRINT BT, Budapest, (1993) 69–81.
- [T46] HARNOS, ZS.: *Sustainability: A system analitic approach Strategies for sustainable agriculture*, MTA Mezőgazdasági Kutatóintézete, Martonvásár, 1993. szeptember 20–26, Conference proceedings, 21–27.
- [T47] HARNOS ZS.: *Az informatika szerepe az agrárkutatásban*, Gazdálkodás, XXXVII. évf. 10. sz., (1993) 65–68.

- [T48] HARNOS ZS.: *Az informatika szerepe az agrárgazdaságban, oktatási célkitűzések*, Informatika a felsőoktatásban, Debreceni Universitas, II. kötet, Debrecen, (1993) 78–84.
- [T49] HARNOS ZS.: *Felkészülés az aszályra - MTA javaslat*, Kertészet és Szőlészet, 42. évf. 3. sz., (1993)
- [T50] HARNOS ZS., SZENTELEKI K.: *Mezőgazdasági és élelmiszeripari információs rendszerek*, Mezőgazdasági Adatbázisok Workshop Kiadványa (1994) 43–50.
- [T51] HARNOS ZS.: *Az agrárgazdálkodást támogató informatikai és döntéstámogató rendszer*, AGRO-21 füzetek 8. szám (1995) 3–14.
- [T52] HARNOS ZS.: *Az ökológiai tájtermelés módszertani eszközei*, AGRO-21 füzetek 8. szám (1995) 15–49.
- [T53] HARNOS ZS. AT AL *Agroökológiai integrált információs rendszer*, AGRO-21 füzetek 8. szám (1995) 50–101.
- [T54] HARNOS, ZS.: *The Climatic Change and the Tendency in Hungary*, EN+IN Conference Proceedings, Budapest, (1995) 7–15.
- [T55] HARNOS ZS., SZENTELEKI K.: *Mezőgazdasági és élelmiszeripari információs rendszerek a Kertészeti és Élelmiszeripari Egyetemen*, Agrárkapcsolatok, II. évf. 2., (1995) 26–27.
- [T56] HARNOS ZS. AT AL: *Az agrárgazdaság fenntartható fejlődésének tudományos megalapozása*, AGRO-21 füzetek 12. szám (1995) 6–116.
- [T57] HARNOS, ZS. AT AL: *On the practice of OR in Hungary*, European Journal of Operation Research, 87 (1995) 452–455.
- [T58] HARNOS ZSOLT: *A kutatás, oktatás és szaktanácsadás szerepe a fenntartható agrárgazdaságban*, A fenntartható fejlődés időszerű kérdései a mezőgazdaságban, II. kötet (1995) 279–284.
- [T59] HARNOS ZSOLT: *BORINFO: Bortermelési, értékesítési és marketing információs rendszer*, Borászati füzetek, VII. évf. 3. sz., (1995) 22–26.
- [T60] HARNOS ZSOLT: *Informatika szerepe az agrártudományokban*, (székfoglaló előadás), MTA Agrártudományok Osztálya Tájékoztatója (1997) 110–117.
- [T61] HARNOS ZS., SZENTELEKI K.: *Az informatika szerepe a minőség biztosításában*, Budapest, (1997), AGRO-21 Füzetek, 13. szám 77–86.
- [T62] HARNOS ZS.: *A klímaváltozás várható alakulása és hatása néhány gazdasági növény termesztettségére* in Az éghajlatváltozás és következményei, Meteorológiai Tudományos Napok '97, OMSz, Budapest, (1998) 55–67.
- [T63] HARNOS ZS., SZENTELEKI K.: *Informatika a szőlő-bor ágazat minőségbiztosításában*, „AGRO-21” Füzetek, 28. szám, (1999) 80–90.

- [T64] HARNOS Zs.: *Informatika a minőségbiztosításban*, „AGRO-21” Füzetek 30. szám (1999) 72–77.
- [T65] HARNOS Zs.: *Informatika az agrár-felsőoktatásban, kutatásban*, Agrárinformatika '99, DATE, Debrecen, (1999) 12–16.
- [T66] HARNOS Zs., SZENTELEKI K.: *A szőlő-bor ágazat információs rendszerei*, Agrárinformatika '99, DATE, Debrecen, (1999) 52–56.
- [T67] HARNOS Zs.: *Informatika az agrártudományokban*, MTA közgyűlési előadások 2000, Millennium az Akadémián IV. kötet 785–793., MTA Budapest, (2001)
- [T68] HARNOS Zs.: *Kockázati tényezők elemzése a növénytermesztésben*, MTA közgyűlési előadások 2000, Millennium az Akadémián IV. kötet 495–511., MTA Budapest, (2001)
- [T69] HARNOS Zs.: *Informatika a precíziós mezőgazdaságban*, EU konform mezőgazdaság és élelmiszerbiztonság 78–83. Debreceni Egyetem, ATC KITE Rt. MTA-DE Földművelési Kutatócsoport, Debrecen (2002)
- [T70] T. E. DOWNING, R. E. BUTTERFIELD, M. BINDI, R. J. BROOKS, T. R. CARTER, R. DELÉCOLLE, Zs. HARNOS, P. A. HARRISON, A. IGLESIAS, M. NEW, S. MOSS, J. E. OLESEN, J. L. ORR, J. PORTER, M. A. SEMENOV AND J. WOLF: *Scaling Methods in Regional Integrated Assessments: From Points Upward and from Global Models Downwards*, Integrated Assessment Vol. 3, Nos. 2-3, (2002) 167–187.
- [T71] HARNOS Zs.: *Mezőgazdasági rendszerek modellezésének néhány problémája*, MTA Agrártudományok osztályának 2002. évi tájékoztatója, MTA Budapest, (2003) 258–263.
- [T72] NÉMETH T., HARNOS Zs., NEMÉNYI M.: *Precíziós növénytermesztés – a hatékonyság növelése és a környezetterhelés csökkentése*, Nemzeti Kutatási és Fejlesztési Programok 4. Program, Nemzeti Kutatási és Technológiai Hivatal Budapest, (2004) 67–77.
- [T73] LÁNG I., HARNOS Zs., JOLÁNKAI M.: *Alkalmazkodási stratégia klímaváltozás esetére : nemzetközi tapasztalatok – hazai lehetőségek*, MAG Kutatás, Fejlesztés és Környezet, VETMA Budapest, (2004. június-július) 28–32.
- [T74] HARNOS Zs.: *A klímaváltozás növénytermelési hatásai*, „AGRO-21” Füzetek, 38. szám, (2005) 38–59.
- [T75] HARNOS Zs.: *A klímaváltozás és lehetséges hatásai a világ mezőgazdaságára*, Magyar Tudomány 2005/7, MTA Budapest, (2005) 827–832.
- [T76] HARNOS Zs.: *A klímaváltozás kockázata - hazai teendők*, VAHAVA zárókonferencia, megjelent DC-n, Budapest, MTA, 2006. március 9.
- [T77] HARNOS Zs., CSETE L., JOLÁNKAI MÁRTON: *A biológiai erőforrások hasznosítása*, Láng István 75 éve, MTA Jubileumi Kiadás (megjelenés alatt)

- [T78] HARNOS ZS., LÁNG I.: *A klímaváltozás lehetséges hatásai az agráriumra*, Agrofórum 18. évf. 10. szám, (2007) 4–9.
- [T79] HARNOS ZS.: *A klímaváltozással összefüggő hazai kutatások: a VAHAVA folytatása*, „KLÍMA-21” Füzetek 49. szám, (2007) 3–15.
- [T80] HARNOS ZS.: *Climate change and some impact on the environment and agriculture modelling case study*, „KLÍMA-21” Füzetek English special edition 55. szám, (2008) 5–23.
- [T81] HARNOS ZS., HUFNAGEL L.: *Kommentár a Nemzeti Éghajlatváltozási Stratégia kialakításához készített természetvédelmi szakértői anyaghoz*, Természetvédelmi Közlemények 14. szám, (2008) 41–44.

**MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA**

KAS PÉTER

(1949–2009)



Kas Péter 1972-ben végezte el az ELTE matematikus szakát. Végzés után az MTA Számítóközpontjába, az MTA SZTAKI egyik jogelődjébe került. Pályáját Prékopa András Operációkutatási osztályán kezdte, amely akkoriban a magyar operációkutatás talán legfontosabb centruma volt.

Itt a Hálózati folyamatok csoportban Klafszyk Emil vezetése alatt dolgozott, akinél a szakdolgozatát írta. Mester és tanítvány soha nem véletlenül találhatnak egymásra. Ebben az esetben összekötötte őket a stílus, a matematizálás igényes, műves volta, mely mindkettejükre halálukig jellemző volt. Előreszaladva az időben, Kas Péter utolsó munkahelyén az Eastern Mediterranean Universityn talán negyven-szer is elmondta a bevezető lineáris algebra tárgyat, mert egyes félévekben több osztályt is vitt. Mégis élete utolsó napjáig nagy gonddal csiszolta az előadásait és készítette a feladatokat a gyakorlatokra és dolgozatokra. A rövidesen induló félév előadásaihoz készített jegyzetei ott feküdtek az asztalán a halálakor.

Első publikációi természetesen a hálózati folyamatok elméletéről és alkalmazásairól szólnak. A témák elég változatosak: a magyar úthálózat forgalmának előrejelzése; egy ütemezési feladat visszavezetése hálózati folyamra; kvadratikusan célfüggvényű folyam feladat; a k -adik legrövidebb út meghatározása; általános nemlineáris, minimál költségű folyam. Ez a korszaka egybeesik a SZTAKI-ban töltött idővel. Ekkor írja egyetemi doktori disszertációját is, ami erőforrások hálózati optimalizálással meghatározott kiosztásáról szól. Az 1993-ban még utoljára visszatér a területhez a tervütemezési feladat kritikus útjainak elemzése során.

1982-ben Prékopa András munkatársaként átment a Műegyetemre. Pályáját egy váratlan fordulat lendítette tovább. 1984 szeptemberében az ELTE Természettudományi Karán megalakult az Operációkutatási Tanszék ugyancsak Prékopa András vezetésével. Kas Péter a tanszék egyik alapító tagja. 1990-ben Ankarába, Törökország akkor egyetlen magánegyetemére, a Bilkent Egyetemre ment, ahol az operációkutatási tanszék szerepét betöltő Department of Industrial Engineeringhez csatlakozott. Innen 1992-ben – hívásra – átment Ciprusra, a famagustai Eastern Mediterranean University (EMU) matematikai tanszékére. Itt tanított 17 tanéven keresztül, egészen a haláláig. Még az 1990-es évek első felében tagja lett annak a bizottságnak, amely megalapította az EMU „industrial engineering” tanszékét. Ő maga azonban maradt a matematikai tanszéken. Itt főképpen lineáris algebrát, valószínűségszámítást, statisztikát, operációkutatást, véges matematikát és különböző optimalizálási tárgyakat tanított. A tanszéknek vannak saját hallgatói is, de legtöbbször más tanszékek számára tanít, ún. szervízkurzusként. Emiatt minden félévben több osztályban indul lineáris algebra, melynek Kas Péter volt a koordinátora, azaz vezető tanára. Ketten doktoráltak és számos szakdolgozatot írtak nála. Szép előadásain kívül tanári sikereinek másik kulcsa az volt, hogy elfogulatlanul közeledett bárkihez, legyen az akár egy tanár kolléga, avagy egy diák.

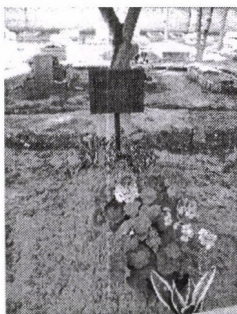
Érdeklődése ciprusi éveiben kezdett a nemlineáris programozás felé fordulni. Vizsgálódásainak homlokterében főleg az entrópia és Young-programozás, valamint a hozzájuk kapcsolódóan divergencia függvények felhasználása állt. Társ szerzőkkel feldolgozott vegyes témákat is.

* * *

Kas Péter 2009. szeptember 14-éről 15-ére virradó éjjel álmában halt meg. Nyugodtan feküdt ágyában, arcán nem látszott szenvedés. Úgy halt meg, ahogy szeretett volna, teljes szellemi és testi erőben. Nem kellett megízlelnie a leépülés keserűségét.

Tanítványainak halála miatt érzett mély, emberei fájdalomából érzékelhető volt az a nagy, pozitív hatás, amit környezetére tett. Ő is azok közé tartozik, akik gyarapították hazánk és a magyar matematika jó hírét.

Sírja Famagusta angol temetőjében található.



Publikációi

- [1] BAKÓ A., KAS P., KIRÁLY L., KLAFSZKY E.: *Operációkutatási módszerek közlekedési hálózatok modellezésére, tervezésére, forgalmi paramétereinek meghatározására*, Kutatási jelentés, KÖTUKI/4/1975, 32.
- [2] KAS, P.: *Kvadrátikus minimálköltségű folyamfeladatról*, Alkalmazott Matematikai Lapok **1**, (1975) 101–108.
- [3] BAKÓ, A., KAS, P.: *A k-adik legrövidebb útvonal meghatározása mátrix módszerrel*, SIGMA **X**, (1977) 60–67.
- [4] KAS, P.: *Megszakítható ütemezési feladatok vizsgálata hálózati folyam mód-szerekkel*, Alkalmazott Matematikai Lapok, **3**, (1977) 131–138.
- [5] KAS, P.: *On a special type job sequencing problem*, Survey of Math. Prog., Proceedings of the 9th Math. Prog. Symp., North-Holland Publ. Comp. **2**, (1979) 279–283.
- [6] KAS, P., MAYER J.: *A nemlineáris folyamprobléma egy megoldási módjáról*, Alkalmazott Matematikai Lapok, **5**, (1979) 157–164.
- [7] KAS, P., MAYER J.: *A reduced gradient approach to the nonlinear network flow problem*, Research Report , MTA SZTAKI, MO/1,1980.
- [8] BAKÓ A., SZIRÁNYI Z., KAS P., BOSZNAI Á.: *Gépjárműjavítás ütemezésé-nek megvalósítása mikrogéppel*, Gyártásügyviteli és Irodagépesítési Országos Tanácskozás (Debrecen) kiadványa, (1985) 11–16.
- [9] (Két társszerzővel) *Épületek rekonstrukciója sorrendjének meghatározása az értékelemzés módszerével*,¹ BME/EPSZERK/8/1985.
- [10] BAKÓ A., BOSZNAY Á., KAS P.: *Leggazdaságosabb útvonalat megadó tájé-koztatási rendszer*, Kutatási zárójelentés, Budapest, 1986.
- [11] BAKÓ A. – KAS P. – KIRÁLY L.: *MPM Optimalizáló programrendszer mik-rogépen*, MCCC, 1987.
- [12] KAS, P., KLAFSZKY, E.: *Megjegyzés egy több dimenziós Cauchy eloszlásról*, Alkalmazott Matematikai Lapok **13**, (1987–88) 145–161.
- [13] BAKÓ A. – KAS P.: *A HDM-III programrendszer felépítése*, KTE Munkabi-zottsági jelentés, CityplaceBudapest, (1989) 30.
- [14] KAS, P., KLAFSZKY, E.: *On the mixed entropy programming*, Research Report, Bilkent University, IE Dept., 1991.
- [15] KAS, P.: *A feasibility theorem on the earliest starting and latest finishing times of critical path analysis* Research Report, Bilkent University, IE Dept., 1991.

¹ Angolból magyarra visszafordított cím.

- [16] KAS, P., KLAFSZKY, E.: *A proof of the generalized Hadamard inequality via information theory (with E. Klafszky)*, Annales Univ. Sci. Budapest, Sect. Comp. **13**, (1992) 21–24.
- [17] KAS, P., KLAFSZKY, E.: *On the dual of the mixed entropy programming problem (with E. Klafszky)*, Optimization **27**, (1993), 253–258.
- [18] KAS, P.: *A feasibility theorem on the earliest starting and latest finishing times of critical path analysis* Central European Journal for Operations Research and Economics **2**, No. 4, (1993) 319–324.
- [19] KAS, P., KLAFSZKY, E.: *On the dual of linear inverse problems (with E. Klafszky)*, European Journal of Operational Research **91**, (1996) 634–639.
- [20] IZBIRAK, G., KAS, P., KLAFSZKY, E., MÁLYUSZ, L.: *Young programming, an analytical approximation of linear programming*, Research Report, Eastern Mediterranean University, AS 97–17.
- [21] KAS, P., KLAFSZKY, E., MÁLYUSZ, L.: *A lineáris programozás analitikus közelítése Young programozással*, in: Komlósi Sándor, Szántai Tamás (szerk.) Új utak a Magyar Operációkutatásban. In memoriam Farkas Gyula., (Válogatás a 23. Magyar operációkutatási konferenciaelőadásából, Dialóg Campus Kiadó, Pécs, (1999)) 144–167.
- [22] KAS, P., KLAFSZKY, E., MÁLYUSZ, L.: *Rácsos tartók állapotvizsgálata*, in: Komlósi Sándor, Szántai Tamás (szerk.) Új utak a Magyar Operációkutatásban. In memoriam Farkas Gyula., (Válogatás a 23. Magyar operációkutatási konferenciaelőadásából, Dialóg Campus Kiadó, Pécs, (1999)) 305–323.
- [23] KAS, P., KLAFSZKY, E., MÁLYUSZ, L.: *Divergence functions in equilibrium problems of trusses*, Publ. Univ. of Miskolc, Series D **39**, (1999) 47–56.
- [24] KAS, P., KLAFSZKY, E., MÁLYUSZ, L.: *Convex programs based on the Young inequality and its relation to linear programming*, Central European Journal for Operations Research and Economics. **7** (2000), 291–304.
- [25] IZBIRAK, G., KAS, P., KLAFSZKY, E., MÁLYUSZ, L.: *Approximation of linear programs by Bregman’s DF projections*, European Journal of Operational Research **126**, (2000) 69–79.
- [26] IZBIRAK, G., KAS, P., KLAFSZKY, E., MÁLYUSZ, L.: *Young Programming*, in: C. A. Floudas and P. M. Pardalos (szerk.), Encyclopedia of Optimization, Kluwer Academic Publisher, Vol. **5**, (2001) 551–557.
- [27] BORATAS, Z., ILLÉS, T., KAS, P.: *Entropy and Young programs: relations and self-concordance*, Central European Journal of Operations Research, **10**, (2002) 261–276.
- [28] KAS, P., KASSAY, G., SENSOY, Z. B.: *On generalized equilibrium points*, Journal of Mathematical Analysis and Applications **296**, (2004) 619–633.

- [29] HAMMER, P. L., KAS, P., VIZVÁRI, B.: *A simple test for integer programming problems with bounded integer variables*, Annales Univ. Sci. Budapest **48**, (2005) 9–14.
- [30] FRENK, J. B. G., KAS, P., KASSAY G.: *On linear programming duality and necessary and sufficient conditions in minimax theory*, Journal of Optimization Theory and Application **132**, (2007) 423–439, <http://dx.doi.org/10.1007/s10957-007-9164-6>

Az Alkalmazott Matematikai Lapok megjelenését támogatja
a Magyar Tudományos Akadémia Könyv- és Folyóiratkiadó Bizottsága.

A kiadásért felelős a BJMT főtítkára
Szedte és tördelte Éliás Mariann

Nyomta a Nagy és Társa Kft., Budapest
Felelős vezető: Földi Gábor

Budapest, 2010
Megjelent 18 (A/5) ív terjedelemben
250 példányban
HU ISSN 0133-3399

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

ÜTMUTATÁS A SZERZŐKNEK

Az Alkalmazott Matematikai Lapok csak magyar nyelvű dolgozatokat közöl. A közlésre szánt dolgozatokat e-mailen az `aml@math.elte.hu` címre kérjük elküldeni az ábrákat tartalmazó fájlokkal együtt. Előnyben részesülnek a \LaTeX -ben elkészített dolgozatok.

A kéziratok szerkezeti felépítésének a következő követelményeket kell kielégíteni:

Fejléc: A fejlécnek tartalmaznia kell a dolgozat címét és a szerző teljes nevét.

Kivonat: A fejléc után egy, képletet nem tartalmazó, legfeljebb 200 szóból álló kivonatot kell minden esetben megadni.

Fejezetek: A dolgozatot címmel ellátott szakaszokra kell bontani, és az egyes szakaszokat arab sorszámozással kell ellátni. Az esetleges bevezetésnek mindig az első szakaszt kell megnevezni.

A dolgozatban előforduló képleteket a dolgozat szakaszokra bontásától független, folytatólagos arab sorszámozással kell azonosítani. Természetesen nem szükséges minden képletet számozással ellátni, csak azokat, amelyekre a szerző a dolgozatban hivatkozni kíván.

Mind az ábrákat, mind a lábjegyzeteket szintén folytatólagos arab sorszámozással kell ellátni. Az ábrák elhelyezését a dolgozat megfelelő helyén ábraazonosító sorszámokkal kell megadni. A lábjegyzetekre a dolgozaton belül az azonosító sorszám felső indexkénti használatával lehet hivatkozni.

Az esetleges definíciókat és tételeket (segédteteleket és lemmákat) szakaszonként újrakezdődő, ponttal elválasztott, kettős számozással kell ellátni. Kérjük a szerzőket, hogy ezeket, valamint a tételek bizonyítását a szövegben kellő módon emeljék ki.

Irodalomjegyzék: A dolgozatok szövegében az irodalmi hivatkozás számait szögletes zárójelben kell megadni, mint például [2] vagy [1, 7–13].

Az irodalmi hivatkozások formája a következő: Minden hivatkozást fel kell sorolni a dolgozat végén található irodalomjegyzékben, a szerzők, illetve a társszerzők esetén az első szerző neve szerint alfabetikus sorrendben úgy, hogy a cirill betűs szerzők nevét a Mathematical Reviews átírási szabályai szerint latin betűsre kell átírni. A folyóiratban megjelent cikkekre [1], a könyvekre [2] a következő minta szerint kell hivatkozni:

[1] FARKAS, J.: *Über die Theorie der einfachen Ungleichungen*, Journal für die reine und angewandte Mathematik 124, (1902) 1–27.

[2] ZOUTENDIJK, G.: *Methods of Feasible Directions*, Elsevier Publishing Company, Amsterdam and New York (1960), 120 o.

Szerző adatai: Az irodalomjegyzék után, a kézirat befejezéseképpen fel kell tüntetni a szerző teljes nevét és a munkahelye (esetleg lakása) pontos címét, illetve e-mail címét.

Idegen nyelvű kivonat: Minden dolgozathoz csatolni kell egy angol nyelvű összefoglalót.

A szerzők a dolgozatukról 20 darab ingyenes különlenyomatot kapnak. A dolgozatok után szerzői díjat az Alkalmazott Matematikai Lapok nem fizet.

317.471

27/2010

TARTALOMJEGYZÉK

| | |
|---|-----|
| <i>Mihálykóné Orbán Éva, Mihálykó Csaba, Lakatos G. Béla</i> , Szintátmetszési probléma és általánosítása a Sparre Andresen-modellben | 1 |
| <i>Kardos Péter, Németh Gábor, Palágyi Kálmán</i> , Bejárásfüggetlen szekvenciális vékonyítás . | 17 |
| <i>Móczár József</i> , A fizikai matematika legújabb eredményei mint a közgazdaság-tudomán lehetséges vizsgálati eszközei | 41 |
| <i>Klafszky Emil</i> , (1934–2009) | 79 |
| <i>Harnos Zsolt</i> , (1941–2009) | 85 |
| <i>Kas Péter</i> , (1949–2009) | 101 |

INDEX

| | |
|---|-----|
| <i>Éva Orbán-Mihálykóné, Csaba Mihálykó, Béla G. Lakatos</i> , Level-crossing problem and its generalization in the Sparre Andresen model | 1 |
| <i>Péter Kardos, Gábor Németh, Kálmán Palágyi</i> , Order-independent sequential thinking | 17 |
| <i>József Móczár</i> , The newest results of physical mathematics as the possible investigation tools of economics | 41 |
| <i>Emil Klafszky</i> , (1934–2009) | 79 |
| <i>Zsolt Harnos</i> , (1941–2009) | 85 |
| <i>Péter Kas</i> , (1949–2009) | 101 |

Alkalmazott matematikai lapok

317.471

13

2010/2

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI TUDOMÁNYOK
OSZTÁLYÁNAK KÖZLEMÉNYEI

27.

KÖTET

ALKALMAZOTT MATEMATIKAI LAPOK

A MAGYAR TUDOMÁNYOS AKADÉMIA MATEMATIKAI TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

ALAPÍTOTTÁK

KALMÁR LÁSZLÓ, TANDORI KÁROLY, PRÉKOPA ANDRÁS, ARATÓ MÁTYÁS

FŐSZERKESZTŐ

PÁLES ZSOLT

FŐSZERKESZTŐ-HELYETTESEK

BENCZÜR ANDRÁS, SZÁNTAI TAMÁS

FELELŐS SZERKESZTŐ

VIZVÁRI BÉLA

TECHNIKAI SZERKESZTŐ

KOVÁCS GERGELY

A SZERKESZTŐBIZOTTSÁG TAGJAI

Arató Mátyás, Csirik János, Csiszár Imre, Demetrovics János, Ésik Zoltán, Frank András, Fritz József, Galántai Aurél, Garay Barna, Gécseg Ferenc, Gerencsér László, Györfi László, Győri István, Hatvani László, Heppes Aladár, Iványi Antal, Járai Antal, Káta Imre, Katona Gyula, Komáromi Éva, Komlósi Sándor, Kovács Margit, Krisztin Tibor, Lovász László, Maros István, Michaletzky György, Pap Gyula, Prékopa András, Recski András, Rónyai Lajos, Schipp Ferenc, Stoyan Gisbert, Seidl László, Tusnády Gábor, Varga László

KÜLSŐ TAGOK:

Csendes Tibor, Fazekas Gábor, Fazekas István, Forgó Ferenc, Friedler Ferenc, Fülöp Zoltán, Kormos János, Maksa Gyula, Racsó Péter, Tallos Péter, Temesi József

27. kötet 2. szám

Szerkesztőség és kiadóhivatal: 1027 Budapest, Fő u. 68.

Az Alkalmazott Matematikai Lapok változó terjedelmű füzetekben jelenik meg, és olyan eredeti tudományos cikkeket publikál, amelyek a gyakorlatban, vagy más tudományokban közvetlenül felhasználható új matematikai eredményt tartalmaznak, illetve már ismert, de színvonalas matematikai apparátus újszerű és jelentős alkalmazását mutatják be. A folyóirat közöl cikk formájában megírt, új tudományos eredménynek számító programokat, és olyan, külföldi folyóiratban már publikált dolgozatokat, amelyek magyar nyelven történő megjelentetése elősegítheti az elért eredmények minél előbbi, széles körű hazai felhasználását. A szerkesztőbizottság bizonyos időnként lehetővé kívánja tenni, hogy a legjobb cikkek nemzetközi folyóiratok különszámaként angol nyelven is megjelenhessenek.

A folyóirat feladata a Magyar Tudományos Akadémia III. (Matematikai) Osztályának munkájára vonatkozó közlemények, könyvismertetések stb. publikálása is.

A kéziratok a főszerkesztőhöz, vagy a szerkesztőbizottság bármely tagjához beküldhetők. A főszerkesztő címe:

Páles Zsolt, főszerkesztő

1027 Budapest, Fő u. 68.

A folyóirat e-mail címe: aml@math.elte.hu

Közlésre el nem fogadott kéziratokat a szerkesztőség lehetőleg visszajuttat a szerzőhöz, de a beküldött kéziratok megőrzéséért vagy továbbításáért felelősséget nem vállal.

Az Alkalmazott Matematikai Lapok előfizetési ára évfolyamonként 1200 forint. Megrendelések a szerkesztőség címén lehetségesek.

A Magyar Tudományos Akadémia III. (Matematikai) Osztálya a következő idegen nyelvű folyóiratokat adja ki:

1. Acta Mathematica Hungarica,
2. Studia Scientiarum Mathematicarum Hungarica.

ZART: EGY MULTIFUNKCIONÁLIS MINTAKERESŐ ALGORITMUS

SZATHMÁRY LÁSZLÓ ÉS BOGNÁR KATALIN

Az adatbányászatban az asszociációs szabályok feltárására számos algoritmust ismerünk. Ebben a cikkben a *Zart* nevű multifunkcionális mintakereső algoritmust mutatjuk be, mely a *Pascal*-algoritmuson alapszik. A *Zart* algoritmus számos olyan műveletet is elvégez, melyek általában egymástól függetlenek, mint például gyakori zárt minták keresése, ill. generátorok hozzárendelése a megfelelő lezártakhoz. A *Zart* ezáltal egy komplett algoritmus a minták ekvivalenciaosztályainak a kiszámítására, melyekben a generátorok és a zárt minták is szerepelnek. Ezen jellemzők miatt a *Zart*-tal könnyen előállíthatók az ún. minimális nem redundáns asszociációs szabályok, melyek az összes asszociációs szabály hasznos és veszteségmentes reprezentációját alkotják. Ezen túlmenően a *Zart* – köszönhetően annak, hogy a *Pascal* ki-terjesztése – elég hatékonyan működik mind gyengén, mind pedig erősen korrelált adathalmazokon. A *Zart* algoritmus kiemelt helyen szerepel a CORON rendszerben, mely egy többcélú adatbányász platform.

1. Bevezetés

Az adatbányászatban az asszociációs szabályok feltárása az egyik legfontosabb feladat. Nagyon sok esetben azonban a gyakori mintákból igen nagy számú érvényes asszociációs szabályt lehet előállítani, erősen megnehezítve ezáltal a szabályok felhasználását egy valódi alkalmazásban. Ezen probléma megoldására megszülettek a szabályok különböző tömörített reprezentációi, mint pl. a generikus bázis (\mathcal{GB}) [1], az informatív bázis (\mathcal{IB}) [1], a reprezentatív szabályok [2], a Duquennes–Guigues-bázis [3], a Luxenburger-bázis [4], stb. Egy nagyon jó összehasonlító elemzés található ezen bázisokról Kryszkiewicz cikkében [5], amelyben a szerző megfogalmazza azokat a feltételeket, melyeknek egy szabály-reprezentációnak meg kell felelnie: legyen *veszteségmentes* (vagyis valamennyi érvényes szabályt vissza lehessen állítani), *helyes* (azaz érvénytelen szabályt ne lehessen belőle előállítani), ill. *informatív* (vagyis bizonyos paramétereket – mint pl. *support* vagy *confidence* értékek – meg lehessen állapítani).

Kryszkiewicz megmutatta, hogy a minimális nem redundáns asszociációs szabályok¹ (\mathcal{MNR}) a *cover* operátorral valamennyi érvényes asszociációs szabály veszteségmentes, helyes, ill. informatív reprezentációját alkotják [5]. Ugyanez érvényes a tranzitív kapcsolatokat nem tartalmazó redukált minimális nem redundáns asszociációs szabályokra (\mathcal{RMNR}) is, ha a *cover* operátort kiegészítjük a *confidence tranzitivitás* jellemzővel. Az \mathcal{MNR} és \mathcal{RMNR} definícióiból látható, hogy

¹Ezen szabályokat a 2. részben definiáljuk.

ezen szabályok a gyakori zárt minták, ill. a zárt mintákhoz rendelt generátorok segítségével előállíthatók. A gyakori mintáknak számos tömörített reprezentációja létezik, mint pl. a zárt minták [6, 7, 8], a generátor reprezentáció [9], a diszjunkciómentes minták [10], a diszjunkciómentes generátorok [9], az általánosított diszjunkciómentes generátorok [11] stb. Ezen reprezentációk közül a gyakori zárt mintákat, ill. a gyakori generátorokat felhasználva lehetőségünk nyílik az asszociációs szabályoknak egy olyan tömörített halmazát definiálni, mely veszteségmentes, helyes, ill. informatív is egyben [5]. Ezt a szabályhalmazt minimális nem redundáns szabályoknak (\mathcal{MNR}) hívjuk [1]. Ez a halmaz – méretét tekintve – általában nem a legkisebb, viszont jó kompromisszumot nyújt a szabályok száma, ill. a szabályok előállításához szükséges idő között [12].

A [13]-as cikkben Bastide *et al.* bemutatta a *Pascal*-algoritmust, s azt állította, hogy ezen algoritmussal előállíthatók az \mathcal{MNR} szabályok. Sajnos a *Pascal*-algoritmus kimenetéből még nem lehet közvetlenül előállítani ezen szabályokat. Először is szükségünk van a gyakori zárt mintákra is. Másodszor, a gyakori generátorokat hozzá kell rendelni a lezártjaikhoz. A *Zart* algoritmus, mely a *Pascal* kiterjesztése, pontosan ezeket a kiegészítő műveleteket végzi el. Az algoritmus neve amúgy a magyar „zárt” szóból származik, ezzel is utalva arra, hogy az algoritmus a gyakori zárt mintákat is megkeresi. A *Zart* algoritmus kimenetéből már könnyedén előállíthatók az \mathcal{MNR} szabályok. A *Pascal* helyett másik algoritmust is vehettünk volna alapul. A választásunk mégis azért esett a *Pascal*-ra, mert a szintenkénti elven működő gyakori mintákat kereső algoritmusok közül talán ez a leghatékonyabb. Ez a jó teljesítmény annak köszönhető, hogy a mintákat számláló következtető mechanizmus jelentős mértékben csökkenti a költséges adathalmaz-újraolvasások számát.

A *Zart* algoritmust a CORON platformon belül implementáltuk [14].² A CORON egy többféle kutatási területen alkalmazható, platformfüggetlen, többcélú adatbányászati eszköztár, amely nem csupán számos adatbányászati algoritmust fog össze, de többféle kiegészítő szolgáltatást is nyújt, mint pl. az adatok előkészítése, tisztítása, ill. az eredmények megjelenítése, értelmezése. Legjobb tudomásunk szerint nem létezik még egy olyan, a CORON-hoz hasonló adatbányászati szoftver, mely direkt módon mintakeresésre, ill. asszociációs szabályok előállítására lett volna kifejlesztve. Cikkünk egy korábban angolul megjelent közleményünk [15] kibővített változata.

Ezen cikk a következőképpen épül fel. A 2. részben áttekintjük az alapvető fogalmakat és definíciókat. A 3. részben a *Zart* algoritmus három fő jellemzőjét mutatjuk be. A 4. részben a *Zart* algoritmust ismertetjük, melynek működését egy konkrét példán keresztül is szemléltetjük. Az 5. részben az \mathcal{MNR} szabályok előállítását részletezzük. A 6. részben a *Zart* algoritmus futási eredményét hasonlítjuk össze a *Pascal*- és *Apriori* algoritmusokkal. Végül a 7. részben a konklúziókkal zárjuk a cikket.

²<http://coron.loria.fr>

2. Fogalmak és definíciók

Vegyük a következő 5×5 -ös méretű adathalmazt: $\mathcal{D} = \{(1, abde), (2, ac), (3, abce), (4, bce), (5, abce)\}$ (lásd még 1. táblázat). A cikk további részében erre a példára úgy fogunk utalni mint „ \mathcal{D} adathalmaz”.

1. táblázat. A példákhoz használt \mathcal{D} adathalmaz.

| | a | b | c | d | e |
|---|---|---|---|---|---|
| 1 | x | x | | x | x |
| 2 | x | | x | | |
| 3 | x | x | x | | x |
| 4 | | x | x | | x |
| 5 | x | x | x | | x |

Gyakori minták. Legyen $\mathcal{O} = \{o_1, o_2, \dots, o_m\}$ objektumok (vagy tranzakciók) egy halmaza, $\mathcal{A} = \{a_1, a_2, \dots, a_n\}$ attribútumoknak egy halmaza, $\mathcal{R} \subseteq \mathcal{O} \times \mathcal{A}$ pedig egy bináris reláció. Az \mathcal{A} halmaz egy részhalmazát *mintának*, egy k darab attribútumot tartalmazó mintát pedig k -hosszúságú mintának nevezünk. Minden tranzakció rendelkezik egy egyedi azonosítóval (*tid*), s a tranzakciók egy halmazát *tidset*nek hívjuk. Egy X mintát tartalmazó tidsetet az X minta *képének* is nevezzük, s $t(X)$ -szel jelöljük. Pl. a \mathcal{D} adathalmazban az ab minta képe 135, azaz $t(ab) = 135$.³ Egy X minta *support* értéke – jelölése $\text{supp}(X)$ – az X minta képének a mérete, azaz $\text{supp}(X) = |t(X)|$. Vagyis a support azt mutatja meg, hogy egy minta hány tranzakcióban van jelen. Egy X mintát gyakran hívunk, ha a support értéke nem kisebb, mint egy adott *minimum support* küszöbérték (jelölése min_supp), azaz $\text{supp}(X) \geq \text{min_supp}$.

Ekvivalenciaosztályok. Két minta $X, Z \subseteq \mathcal{A}$ ekvivalens ($X \cong Z$), ha a képük megegyezik, vagyis $t(X) = t(Z)$. Egy P mintával ekvivalens minták halmazát a P minta ekvivalenciaosztályának nevezzük: $[P] = \{Q \subseteq \mathcal{A} \mid P \cong Q\}$. Egy ekvivalenciaosztálynak egyetlen maximális eleme van (zárt minta), s több minimális eleme is lehet (generátorok⁴) [13].

2.1. Definíció. Egy minta *zárt*, ha nincs vele azonos support értékű valódi szuperhalmaza.

³A halmazok leírására egyszerűsített jelölést alkalmazunk. Pl. $\{a, b, e\}$ helyet abe -t, $\{1, 3, 5\}$ helyett pedig 135-öt fogunk írni.

⁴A szakirodalomban ezen mintákat többféleképpen is nevezik: kulcsminták, minimális generátorok, szabad minták, kulcsgenerátorok stb.

2.2. Definíció. Egy minta (minimális) *generátor*, ha nincs vele azonos support értékű valódi részhalmaza.

A *lezárási* operátor egy X mintához az X ekvivalenciaosztályában található maximális elemet rendeli hozzá. A hozzárendelt elemet X *lezártjának* nevezzük, s $\gamma(X)$ -szel jelöljük.

1. Példa. A \mathcal{D} adathalmazban zárt minta pl. az a , ac és $abce$. Az abc minta (support értéke 2) nem zárt az $abce$ minta miatt (melynek support értéke szintén 2). Az abc minta generátor, ui. valamennyi részhalmaza magasabb support értékkel rendelkezik. A b és abc minták generátorok, melyek lezártjai: $\gamma(b) = be$ és $\gamma(abc) = abce$.

Mintaszámláló következtetés. A support értékre és a generátorokra vonatkozó most következő tulajdonságok a [13]-as cikkből származnak:

1. Tulajdonság. Legyen P és Q egy-egy minta.

- (i) $P \cong Q \Rightarrow \text{supp}(P) = \text{supp}(Q)$,
- (ii) $P \subseteq Q$ és $(\text{supp}(P) = \text{supp}(Q)) \Rightarrow P \cong Q$.

2. Tulajdonság. Egy gyakori generátor valamennyi részhalmaza gyakori generátor.

3. Tulajdonság. Egy P minta a.cs.a. generátor, ha $\text{supp}(P) \neq \min_{p \in P} (\text{supp}(P \setminus \{p\}))$.

Az utolsó jellemző szerint egy P minta a.cs.a. generátor, ha P support értéke különbözik a nála eggyel kisebb méretű részhalmazainak a support értékétől. Definíció szerint ui. egy generátornak nem lehet vele azonos support értékű valódi részhalmaza.

Gyakori asszociációs szabályok. Egy asszociációs szabály egy $I_1 \rightarrow I_2$ alakú kifejezés, ahol I_1 és I_2 tetszőleges minták ($I_1, I_2 \subseteq \mathcal{A}$), $I_1 \cap I_2 = \emptyset$ és $I_2 \neq \emptyset$. Egy $r: I_1 \rightarrow I_2$ asszociációs szabály support értéke a következőképpen definiálható: $\text{supp}(r) = \text{supp}(I_1 \cup I_2)$. Egy r asszociációs szabály *confidence* értéke azt mutatja meg, hogy mi annak a valószínűsége, hogy egy objektum rendelkezik az I_2 mintával, feltéve hogy rendelkezik az I_1 mintával: $\text{conf}(r) = \text{supp}(I_1 \cup I_2) / \text{supp}(I_1)$. Egy r asszociációs szabályt melynek confidence értéke 100%, pontos szabálynak (vagy implikációnak), különben pedig megközelítő szabálynak nevezzük. A supporthoz hasonlóan a confidence-re is megadható egy határérték (*minimum confidence*). Egy r asszociációs szabály érvényes, ha $\text{supp}(r) \geq \min_supp$ és $\text{conf}(r) \geq \min_conf$. Az érvényes asszociációs szabályok halmazát \mathcal{AR} -rel jelöljük.

2. Példa. Nézzünk néhány asszociációs szabályt, melyeket a \mathcal{D} adathalmazból nyerhetünk ki. A $b \rightarrow e$ support értéke 4 (80%), confidence értéke 1 (100%). A $c \rightarrow abe$ support értéke 2 (40%), confidence értéke 0,5 (50%).

Most pedig tekintsük át néhány, asszociációs szabályokat reprezentáló bázis definícióját [16].

2.3. Definíció. Jelölje FCI a gyakori zárt minták halmazát. Egy f gyakori zárt minta esetén jelölje FG_f az f ekvivalenciaosztályában lévő generátorokat. A pontos asszociációs szabályok (implikációk) generikus bázisa:

$$\mathcal{GB} = \{r : g \Rightarrow (f \setminus g) \mid f \in FCI \wedge g \in FG_f \wedge g \neq f\}.$$

2.4. Definíció. Jelölje FCI a gyakori zárt minták, FG pedig a gyakori generátorok halmazát. A megközelítő asszociációs szabályok informatív bázisa:

$$\mathcal{IB} = \{r : g \rightarrow (f \setminus g) \mid f \in FCI \wedge g \in FG \wedge \gamma(g) \subset f\}.$$

2.5. Definíció. Jelölje FCI a gyakori zárt minták halmazát, \mathcal{IB} pedig az előbb definiált megközelítő asszociációs szabályok informatív bázisát. Az \mathcal{IB} tranzitív kapcsolatokról mentes változatát redukált informatív bázisnak nevezzük:

$$\begin{aligned} \mathcal{RIB} = \{r : g \rightarrow (f \setminus g) \in \mathcal{IB} \mid \gamma(g) \\ \text{az } f \text{ legnagyobb valódi részhalmaza az } FCI\text{-ben}\}. \end{aligned}$$

2.6. Definíció. A minimális nem redundáns asszociációs szabályokat a \mathcal{GB} és \mathcal{IB} halmazok uniója adja ($\mathcal{MNR} = \mathcal{GB} \cup \mathcal{IB}$). Az \mathcal{MNR} tranzitív kapcsolatokról mentes változatát a \mathcal{GB} és \mathcal{RIB} halmazok uniójaként kapjuk ($\mathcal{RMNR} = \mathcal{GB} \cup \mathcal{RIB}$).

3. Példa. Nézzünk néhány asszociációs szabályt a fenti bázisokból: $\{b \rightarrow e, ab \rightarrow e, ce \rightarrow b\} \in \mathcal{GB}$, $\{b \rightarrow ce, b \rightarrow ace\} \in \mathcal{IB}$, $\{b \rightarrow ce, b \rightarrow ae, bc \rightarrow ae\} \in \mathcal{RIB}$. Az 5. rész végén további példák is találhatók.

3. A Zart fő jellemzői

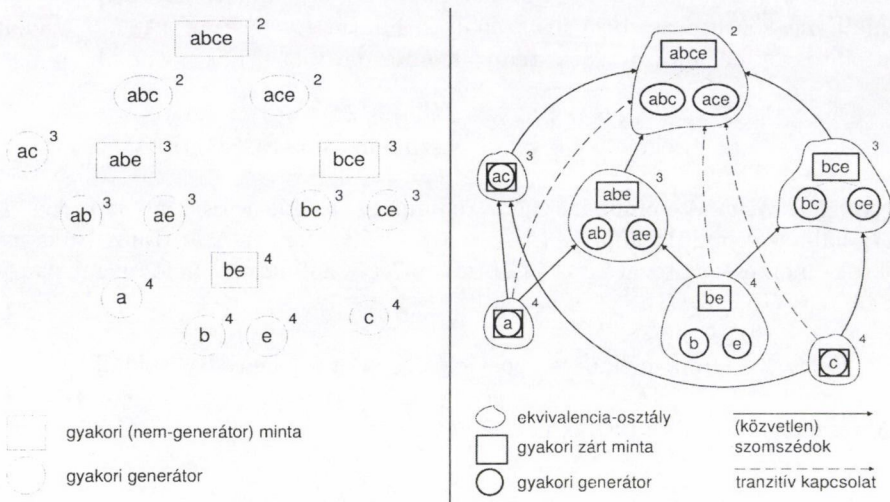
A Zart algoritmus három fázisból tevődik össze: (1) mintaszámláló következtetés, (2) gyakori zárt minták azonosítása, (3) a gyakori zárt minták generátorainak a megállapítása.

3.1. Mintaszámláló következtetés

A Zart első része a Pascal-algoritmus mintaszámláló következtetésére épül. A gyakori minták szintenkénti keresése esetén először az ekvivalenciaosztályok legkisebb elemeit találjuk meg, s ezek a minták pontosan a generátorok. Később, mikor egy nagyobb mintát találunk, leteszteljük, hogy egy már felfedezett ekvivalenciaosztályhoz tartozik-e. Ha igen, akkor nem szükséges újra végigfutni az adathalmazon a minta support értékének megállapításához. Vagyis ezáltal a költséges adathalmaz-újraolvasásokat és support számításokat le lehet csökkenteni csupán a

generátorok esetére. Egy bizonyos szinttől viszont már meglesz az összes generátor, tehát a hátralévő gyakori minták, ill. ezek support értékei megállapíthatók anélkül, hogy újra végig kellene olvasni az adathalmazt.

Az 1. ábrán (bal oldal) a *Pascal* kimenete látható: az algoritmus megtalálja a gyakori mintákat és megjelöli a gyakori generátorokat. Ha megnézzük az MNR és $RMNR$ szabályok definícióit, akkor egyértelmű, hogy ez a kimenet nem elegendő. A *Zart* kimenete szintén az 1. ábrán szerepel (jobb oldal). Mint látható, az algoritmus feltárja az ekvivalenciaosztályokat. Az ábrán csupán az ekvivalenciaosztályok maximális (gyakori zárt minta) és minimális (gyakori generátorok) elemei vannak feltüntetve. A support értékek az osztályok jobb felső sarkában szerepelnek. A *Zart* által produkált kimenet szükséges és elegendő a GB , IB , RIB , MNR és $RMNR$ szabályok generálásához.



1. ábra. A *Pascal* (bal oldal) és a *Zart* (jobb oldal) eredménye a D adathalmazon $min_supp = 2$ (40%) küszöbérték mellett.

3.2. Zárt minták azonosítása a gyakori minták között

A *Zart* algoritmus második része a zárt mintákat azonosítja a gyakori minták között. Ez az ötlet eredetileg az *Apriori-Close* algoritmusban szerepelt [16]. Definíció szerint egy zárt mintának nincs vele azonos support értékű szuperhalmaza. A k . iteráció során valamennyi k -hosszúságú mintát „zárt”-nak minősítünk. A $(k + 1)$. iterációban minden $(k + 1)$ -hosszúságú minta esetén megnézzük, hogy az adott minta tartalmaz-e vele azonos support értékű k -hosszúságú mintát. Ha igen, akkor a k -hosszúságú minta nem zárt, hiszen van egy vele azonos support

értékű superhalmaza. Így a k -hosszúságú minta státusza „nem zárt”-ra változik. Mikor az algoritmus befejezi a futását, a még mindig „zárt”-nak jelölt minták lesznek a tényleges gyakori zárt minták.

3.3. Generátorok és lezártjaik összerendelése

A szintenkénti bejárás következtében mikor megtalálunk egy zárt mintát, akkorra már valamennyi gyakori részhalmazát feltártuk. Ez azt jelenti, hogy már a generátorait is érintettük, csupán be kell őket azonosítani. Megmutatjuk, hogy a generátorok keresési tere leszűkíthető a nem zárt mintákra. Ez a következő két tulajdonságon alapszik:

4. *Tulajdonság.* Egy zárt minta nem lehet egy nagyobb minta generátora.

5. *Tulajdonság.* Egy gyakori nem zárt g generátor lezártja nem más, mint a g legkisebb valódi superhalmaza a gyakori zárt minták között.

Ezen két tulajdonságot felhasználva a generátorokat a következőképpen találhatjuk meg hatékonyan. A generátorokat egy l listában tároljuk. A k . iteráció során a k -hosszúságú gyakori zárt mintákat kiszűrjük. Minden k -hosszúságú gyakori zárt minta esetén (nevezzük a mintát z -nek) a következő lépéseket kell elvégezni: keressük meg z részhalmazait az l listában, regisztráljuk őket mint z generátorait, majd töröljük őket l -ből. Mielőtt továbblépnénk a $(k + 1)$. iterációra, adjuk hozzá az l listához a k -hosszúságú nem zárt generátorokat. A 4-es és 5-ös tulajdonságok garantálni fogják, hogy mikor egy gyakori zárt minta részhalmazait keressük az l listában, akkor csakis a generátorait fogjuk megtalálni. A listából visszaadott minták support értéke azonos a gyakori zárt minta supportjával, így ezt nem kell külön leellenőrizni. Mivel a listában csupán a generátorokat tároljuk, így jóval kevesebb elemet kell tesztelni, mintha az összes gyakori mintát letárolnánk. Mivel a k . lépés során az l listában tárolt legnagyobb minta mérete legfeljebb $(k - 1)$ lehet, ezért nem fogjuk megtalálni azon generátorokat, melyek azonosak a lezártjukkal. Ha egy gyakori zárt mintának az algoritmus lefutása után nincs beregisztrált generátora, akkor ez egyszerűen azt jelenti, hogy a minta generátora önmaga. Ami az implementációt illeti, a generátorok tárolására egy hagyományos lista helyett a prefix fa (*trie*) adatstruktúrát javasoljuk, mivel ezen adatstruktúrával nagyon gyorsan meg lehet keresni egy minta részhalmazait.

2. táblázat. A Zart által használt táblák és mezők.

| | |
|-------|---|
| C_i | potenciálisan gyakori i -hosszúságú jelöltek mezők: (1) minta, (2) pred_supp, (3) kulcs, (4) support |
| F_i | gyakori i -hosszúságú minták mezők: (1) minta, (2) kulcs, (3) support, (4) zárt |
| Z_i | gyakori i -hosszúságú zárt minták mezők: (1) minta, (2) support, (3) gen |

| | |
|-----------|---|
| minta | - egy tetszőleges minta |
| pred_supp | - az $(i - 1)$ -hosszúságú gyakori részhalmazok support értékeinek a minimuma |
| kulcs | - az adott minta generátor? |
| zárt | - az adott minta zárt? |
| gen | - az adott zárt minta generátorai |

3. táblázat. A *Zart* algoritmus főblokkja.

- 1) $FG \leftarrow \{ \};$ // globális lista a gyakori generátoroknak
- 2) C_1 feltöltése 1-hosszúságú mintákkal;
- 3) C_1 -ben lévő minták support értékeinek a megállapítása;
- 4) gyakori minták átmásolása C_1 -ből F_1 -be;
- 5) az F_1 -ben lévő mintákat jelöljük meg „zárt”-ként;
- 6) az F_1 -ben lévő mintákat jelöljük meg „kulcs”-ként ha a support értékük kisebb, mint az adathalmaz objektumainak a száma;
// ui. definíció szerint az üres halmaz support értéke 100%
- 7) ha a bemeneti adathalmaznak van telített oszlopa, akkor $FG \leftarrow \{\emptyset\}$;
- 8) $i \leftarrow 1$;
- 9) ciklus
- 10) {
- 11) $C_{i+1} \leftarrow \text{Zart-Gen}(F_i)$;
- 12) ha C_{i+1} üres, akkor lépünk ki a ciklusból;
- 13) állapítsuk meg a „kulcs”-ként megjelölt minták support-ját a C_{i+1} táblában;
- 14) ha a C_{i+1} -ben van olyan c minta, hogy $c.\text{support} = c.\text{pred_supp}$,
akkor jelöljük meg a c mintát „nem kulcs”-ként;
- 15) gyakori minták másolása C_{i+1} -ből F_{i+1} -be;
- 16) ha egy F_{i+1} -ben lévő mintának van F_i -beli részhalmaza úgy, hogy
a két minta support-ja azonos, akkor a részhalmazt jelöljük
meg „nem zárt”-ként;
- 17) a „zárt”-ként megjelölt mintákat másoljuk át F_i -ből Z_i -be;
- 18) **Find-Generators**(Z_i);
- 19) $i \leftarrow i + 1$;
- 20) }
- 21) minták átmásolása F_i -ből Z_i -be;
- 22) **Find-Generators**(Z_i);

4. A Zart algoritmus

4.1. Pszeudó kód

A *Zart* algoritmus három különböző táblát használ, ezek leírása az 2. táblázatban látható. Az algoritmus főblokkja a 3. táblázatban szerepel. Az algoritmus részletesebb pszeudó kódja a [14]-es dolgozatban szerepel.

Zart-Gen függvény. A függvény bemenete egy F_i tábla, mely i -hosszúságú gyakori mintákat tartalmaz. A függvény kimenete a C_{i+1} tábla. Az eljárás a következő: a függvény feltölti a C_{i+1} táblát az F_i -ben szereplő minták 1-gyel nagyobb szuperhalmazával. A C_{i+1} táblában szereplő minták pred_supp értékei a náluk 1-gyel kisebb méretű gyakori részhalmazaik support értékeinek a minimumát fogják felvenni. Ha valamelyik részhalmaz nem generátor, akkor az adott $(i+1)$ -hosszúságú minta sem generátor, így a support értéke azonos lesz a pred_supp értékkel.

Find-Generators eljárás. Az eljárás bemenetként kap egy Z_i táblát, majd a táblában található valamennyi z gyakori zárt minta esetén a következő műveleteket végzi el: megkeresi z valódi részhalmazait a globális FG listában, a találatokat beregisztrálja mint z generátorait, kitörli őket az FG listából, majd az F_i -ben található nem zárt generátorokat hozzáadja az FG listához.

4.2. Futási példa

A *Zart* végrehajtását a \mathcal{D} adathalmazon a 4. táblázatban szemléltetjük. Az algoritmus először végigolvassa az adathalmazt az 1-hosszúságú minták support értékeinek megállapítása végett. A d minta törlésre kerül, mivel nem gyakori. A következő iteráció során valamennyi 2-hosszúságú jelöltet előállítjuk, ill. megállapítjuk a support értékeiket. A C_2 -ben van egy minta, amelynek a support értéke azonos az egyik részhalmazának a supportjával, így a be minta nem generátor. Az F_2 táblázat segítségével megállapítjuk, hogy az F_1 -ben található b és e minták nem zártak, mivel van azonos support-ú szuperhalmazuk. A megmaradt a és c zárt mintákat átmásoljuk a Z_1 táblába, majd megállapítjuk a generátorainkat. A gyakori generátorok globális listájában (FG) – mely még üres – nincs részhalmazuk, ami azt jelenti, hogy az a és c minták generátorok. Az F_1 -ben lévő nem zárt generátorokat (b és e) hozzáadjuk az FG listához. A C_3 táblában az abe és bce minták nem generátorok. Az F_3 tábla alapján az ab , ae , bc és ce minták nem zártak. A megmaradt ac és be zárt mintákat átmásoljuk a Z_2 táblába. Az ac generátora önmaga, míg a be generátorai b és e . Ezen két generátort kitöröljük az FG listából, majd az ab , ae , bc és ce mintákat hozzáadjuk az FG -hez. Az $abce$ jelölt szintén nem generátor, s mivel nincs több generátor-jelölt C_4 -ben, ezért innentől már nem kell többször végigolvasni az adathalmazt a support értékek megállapításához. Az ötödik iteráció során már nem tudunk új jelöltet előállítani, így az algoritmus kilép a ciklusból. Az $abce$ generátorait kiolvassuk az FG listából.

Az algoritmus tehát megkeresi valamennyi gyakori mintát, gyakori zárt mintát, ill. a zárt minták generátorait is (lásd 5. táblázat, jobb oldal). A táblázatban a

4. táblázat. A *Zart* végrehajtása a \mathcal{D} adathalmazon ($\min_supp = 2$ (40%)).

adathalmaz
végigolvasása₁

→

| C_1 | pred_supp | kulcs | supp |
|-------|-----------|-------|------|
| a | | | 4 |
| b | | | 4 |
| c | | | 4 |
| d | | | 1 |
| e | | | 4 |

| F_1 | kulcs | supp | zárt |
|-------|-------|------|------|
| a | igen | 4 | igen |
| b | igen | 4 | igen |
| c | igen | 4 | igen |
| e | igen | 4 | igen |

| Z_1 | supp | gen |
|-------|------|-----|
| a | 4 | |
| c | 4 | |

$FG_{elbűt} = \{\}$

$FG_{utána} = \{b, c\}$

adathalmaz
végigolvasása₂

→

| C_2 | pred_supp | kulcs | supp |
|-------|-----------|-------|------|
| ab | 4 | igen | 3 |
| ac | 4 | igen | 3 |
| ae | 4 | igen | 3 |
| bc | 4 | igen | 3 |
| be | 4 | igen | 4 |
| ce | 4 | igen | 3 |

| F_2 | kulcs | supp | zárt |
|-------|-------|------|------|
| ab | igen | 3 | igen |
| ac | igen | 3 | igen |
| ae | igen | 3 | igen |
| bc | igen | 3 | igen |
| be | nem | 4 | igen |
| ce | igen | 3 | igen |

| Z_2 | supp | gen |
|-------|------|------------|
| ac | 3 | |
| bc | 4 | $\{b, c\}$ |

$FG_{elbűt} = \{b, c\}$

$FG_{utána} = \{ab, ac, bc, ce\}$

adathalmaz
végigolvasása₃

→

| C_3 | pred_supp | kulcs | supp |
|-------|-----------|-------|------|
| abc | 3 | igen | 2 |
| abe | 3 | igen | 3 |
| ace | 3 | igen | 2 |
| bce | 3 | igen | 3 |

| F_3 | kulcs | supp | zárt |
|-------|-------|------|------|
| abc | igen | 2 | igen |
| abe | nem | 3 | igen |
| ace | igen | 2 | igen |
| bce | nem | 3 | igen |

| Z_3 | supp | gen |
|-------|------|--------------|
| abc | 3 | $\{ab, ac\}$ |
| bce | 3 | $\{bc, ce\}$ |

$FG_{elbűt} = \{ab, ac, bc, ce\}$

$FG_{utána} = \{abc, ace\}$

| C_4 | pred_supp | kulcs | supp |
|-------|-----------|-------|------|
| abce | 2 | igen | 2 |

| F_4 | kulcs | supp | zárt |
|-------|-------|------|------|
| abce | nem | 2 | igen |

| Z_4 | supp | gen |
|-------|------|----------------|
| abce | 2 | $\{abc, ace\}$ |

$FG_{elbűt} = \{abc, ace\}$

$FG_{utána} = \{\}$

| C_5 | pred_supp | kulcs | supp |
|-------------|-----------|-------|------|
| \emptyset | | | |

zárt mintákat „+” jellel láttuk el. A support értékeket a minták mellett zárójelben tüntettük fel. Ha az algoritmus üresen hagyja egy zárt minta generátorait, akkor ez azt jelenti, hogy a generátor azonos a zárt mintával (lásd az *a*, *c* és *ac* mintákat a 4. táblázatban).

5. Minimális nem redundáns asszociációs szabályok előállítás a *Zart* algoritmussal

Ha a gyakori mintákból állítjuk elő az összes érvényes asszociációs szabályt, akkor valószínűleg túl sok szabályt fogunk kapni, melyek között számos szabály redundáns. Pl. a \mathcal{D} adathalmazból $\min_supp = 2$ (40%) és $\min_conf = 50\%$ mellett nem kevesebb, mint 50 szabályt tudunk kinyerni. Figyelembe véve az adathalmaz 5×5 -ös méretét, ez a mennyiség óriási. Hogyan tudnánk megtalálni a legérdekesebb szabályokat? Hogyan lehetne elkerülni a redundanciát, ill. csökkenteni a szabályok számát? A minimális nem redundáns asszociációs szabályok (\mathcal{MNR}) a segítségünkre lehetnek. A 2.6. definíció alapján egy \mathcal{MNR} szabály

5. táblázat. A Zart algoritmus kimenete.

| Összes gyakori minta ($\bigcup_i F_i$) | | Összes gyakori zárt minta és generátorai ($\bigcup_i Z_i$) |
|--|-------------|--|
| $a(4) +$ | $be(4) +$ | $a(4); [a]$ |
| $b(4)$ | $ce(3)$ | $c(4); [c]$ |
| $c(4) +$ | $abc(2)$ | $ac(3); [ac]$ |
| $e(4)$ | $abe(3) +$ | $be(4); [b, e]$ |
| $ab(3)$ | $ace(2)$ | $abe(3); [ab, ae]$ |
| $ac(3) +$ | $bce(3) +$ | $bce(3); [bc, ce]$ |
| $ae(3)$ | $abce(2) +$ | $abce(2); [abc, ace]$ |
| $bc(3)$ | | |

formája a következő: a bal oldalon egy gyakori generátor áll, a bal és jobb oldal uniója egy gyakori zárt minta, s a generátor valódi részhalmaza ezen zárt mintának. Az \mathcal{MNR} szabályok előállítására a gyakori zárt mintákra és ezen minták hozzátársított generátoraira van szükség. Mivel a Zart algoritmus mindkét halmazt megkeresi, ezért az algoritmus kimenetéből (lásd 5. táblázat) közvetlenül generálhatók ezek a szabályok. A generátorok valódi zárt szuperhalmazainak a gyors azonosítására érdemes a gyakori zárt mintákat egy prefix fában (*trie*) tárolni.

Az \mathcal{MNR} szabályokat a következőképpen lehet előállítani: minden egyes P_1 gyakori generátornak keressük meg a zárt szuperhalmazait (P_2), majd adjuk hozzá az \mathcal{MNR} halmazhoz az $r: P_1 \rightarrow P_2 \setminus P_1$ szabályt. Pl. az e generátor felhasználásával az 1. ábrán (jobb oldal) három szabály generálható. Az ekvivalenciaosztályokon belül előállított szabályok alkotják a generikus bázist (\mathcal{GB}), melyek mindegyike pontos asszociációs szabály ($e \Rightarrow b$). A különböző ekvivalenciaosztályok között előállított szabályok mind megközelítő szabályok ($e \rightarrow bc$ és $e \rightarrow abc$). Az \mathcal{RMNR} szabályok esetén a generátorokhoz csupán azon gyakori zárt szuperhalmazokat keressük meg, ahol a zárt minta ekvivalenciaosztálya közvetlen szomszédja a generátor ekvivalenciaosztályának. Vagyis, maradva az előző példánál, csupán három \mathcal{RMNR} szabályt tudunk előállítani az e generátorral: $e \Rightarrow b$, $e \rightarrow bc$ és $e \rightarrow ab$. Az 6. táblázatban megfigyelhetők a különböző szabályhalmazok méretei.⁵ Mint látható, ritka adathalmazok esetén (pl. T20I6D100K) az \mathcal{MNR} szabályok száma majdnem ugyanannyi, mint az összes lehetséges szabály száma (\mathcal{AR}). Viszont sűrű, erősen korrelált adathalmazok esetén (pl. C20D10K vagy MUSHROOMS) a szabályhalmazok méretbeli különbsége már jelentős. Az \mathcal{RMNR} halmaz mérete mindig jóval kisebb, mint az \mathcal{AR} halmaz mérete, mind ritka, mind pedig sűrű adathalmazok esetén.

6. táblázat. Asszociációs szabályhalmazok méretbeli összehasonlítása.

| adathalmaz (min_supp) | min_conf | AR (összes érvényes szabály) | GB | IB | RIB | MNR ($GB \cup IB$) | $RMNR$ ($GB \cup RIB$) |
|--------------------------|----------|--------------------------------------|------|-----------|---------|---------------------------|-----------------------------|
| D (40%) | 50% | 50 | 8 | 17 | 13 | 25 | 21 |
| T20I6D100K (0.5%) | 90% | 752,715 | 232 | 721,716 | 91,422 | 721,948 | 91,654 |
| | 70% | 986,058 | | 951,340 | 98,097 | 951,572 | 98,329 |
| | 50% | 1,076,555 | | 1,039,343 | 101,360 | 1,039,575 | 101,592 |
| | 30% | 1,107,258 | | 1,068,371 | 102,980 | 1,068,603 | 103,212 |
| C20D10K (30%) | 90% | 140,651 | 967 | 8,254 | 2,784 | 9,221 | 3,751 |
| | 70% | 248,105 | | 18,899 | 3,682 | 19,866 | 4,649 |
| | 50% | 297,741 | | 24,558 | 3,789 | 25,525 | 4,756 |
| | 30% | 386,252 | | 30,808 | 4,073 | 31,775 | 5,040 |
| MUSHROOMS (30%) | 90% | 20,453 | 544 | 952 | 682 | 1,496 | 1,226 |
| | 70% | 45,147 | | 2,961 | 1,221 | 3,505 | 1,765 |
| | 50% | 64,179 | | 4,682 | 1,481 | 5,226 | 2,025 |
| | 30% | 78,888 | | 6,571 | 1,578 | 7,115 | 2,122 |

6. Teszteredmények

Tesztjeink során a *Zart* algoritmust az *Apriori* és a *Pascal*-algoritmusokkal mértük össze. Mindhárom algoritmust Java-ban implementáltuk a CORON adatbányász platformban [14].⁶ A teszteket egy Intel Pentium IV 2.4 GHz-es gépen végeztük el Debian GNU/Linux operációs rendszer alatt. Az adott gép 512 MB RAM-mal rendelkezett. Valamennyi válaszidő valós idő, melyet a Unix rendszereken használt *time* paranccsal mértünk le a bemenet és a kimenet között. A tesztekhez a következő adathalmazokat használtuk fel: T20I6D100K, C20D10K és MUSHROOMS. A T20⁷ egy ritka adathalmaz, s felépítésében a bevásárlóközpontok adatbázisaira hasonlít, melyek általában gyengén korrelált adatokat tartalmaznak. A C20 egy népszámlálás részadatait tartalmazza, míg a MUSHROOMS különböző gombok jellemzőit írja le. Ez utóbbi két adathalmaz erősen korrelált. Korábbi kutatási eredmények szerint a gyengén korrelált adatok, mint pl. a szintetikus adatok, viszonylag könnyű feladatot jelentenek a gyakori mintákat kereső algoritmusok számára, ui. ezen adathalmazokban kevés a gyakori minta. Az ilyen típusú adatokra valamennyi algoritmus hasonló futási időt produkált. Ezzel szemben viszont a sűrű és erősen korrelált adatok sokkal nagyobb kihívást jelentenek, mely arra vezethető vissza, hogy itt meglehetősen nagy lehet a különbség a gyakori és a gyakori zárt minták száma között. Nagyon sok valós életből vett adathalmaz ilyen tulajdonságokkal rendelkezik. Az algoritmusok által produkált futási eredményeket a 7. táblázatban foglaltuk össze.

⁵ Definíció szerint a GB halmazba tartozó szabályok confidence értéke 100%.

⁶ <http://coron.loria.fr>

⁷ <http://www.almaden.ibm.com/software/quest/Resources/>

7. táblázat. A *Zart* futási ideje, ill. egyéb statisztikák (gyakori minták (GyM-k) száma, gyakori zárt minták (GyZM-k) száma, gyakori generátorok (GyG-ok) száma, GyZM-k és GyM-k aránya, GyG-ok és GyM-k aránya).

| min_supp | futási idő (mp.) | | | # GyM | # GyZM | # GyG | #GyZM #GyM | #GyG #GyM |
|-------------------|------------------|--------|--------|---------|---------|---------|---------------|--------------|
| | Apriori | Pascal | Zart | | | | | |
| T20I6D100K | | | | | | | | |
| 2% | 72.67 | 71.15 | 71.16 | 378 | 378 | 378 | 100.00% | 100.00% |
| 1% | 107.63 | 106.24 | 107.69 | 1,534 | 1,534 | 1,534 | 100.00% | 100.00% |
| 0.75% | 134.49 | 132.00 | 133.00 | 4,710 | 4,710 | 4,710 | 100.00% | 100.00% |
| 0.5% | 236.10 | 228.37 | 230.17 | 26,836 | 26,208 | 26,305 | 97.66% | 98.02% |
| 0.25% | 581.11 | 562.47 | 577.69 | 155,163 | 149,217 | 149,447 | 96.17% | 96.32% |
| C20D10K | | | | | | | | |
| 50% | 61.18 | 16.68 | 17.94 | 1,823 | 456 | 456 | 25.01% | 25.01% |
| 40% | 71.60 | 19.10 | 19.22 | 2,175 | 544 | 544 | 25.01% | 25.01% |
| 30% | 123.57 | 26.74 | 26.88 | 5,319 | 951 | 967 | 17.88% | 18.18% |
| 20% | 334.87 | 53.28 | 54.13 | 20,239 | 2,519 | 2,671 | 12.45% | 13.20% |
| 10% | 844.44 | 110.78 | 118.09 | 89,883 | 8,777 | 9,331 | 9.76% | 10.38% |
| MUSHROOMS | | | | | | | | |
| 60% | 3.10 | 2.04 | 2.05 | 51 | 19 | 21 | 37.25% | 41.18% |
| 50% | 6.03 | 3.13 | 3.13 | 163 | 45 | 53 | 27.61% | 32.52% |
| 40% | 13.93 | 6.00 | 6.03 | 505 | 124 | 153 | 24.55% | 30.30% |
| 30% | 46.18 | 12.79 | 12.84 | 2,587 | 425 | 544 | 16.43% | 21.03% |
| 20% | 554.95 | 30.30 | 34.88 | 53,337 | 1,169 | 1,704 | 2.19% | 3.19% |

6.1. Gyengén korrelált adatok

A T20 szintetikus adathalmaz a bevásárlóközpontok adatbázisait utánozza, így egy ritka és gyengén korrelált adathalmazról van szó. Ezen adathalmazban kevés a gyakori minta, s majdnem valamennyi gyakori minta generátor. Az *Apriori*, *Pascal*- és *Zart* algoritmusok hasonlóképpen viselkednek. Mint látható, a T20 adathalmazban 0.75% minimum support felett valamennyi gyakori minta zárt és generátor is egyszerre. Ez azt jelenti, hogy minden egyes ekvivalenciaosztálynak csupán egyetlen eleme van. Emiatt a *Zart* és *Pascal*-algoritmusok nem tudják kihasználni a mintaszámláló következtetés előnyeit, így ugyanúgy kénytelenek dolgozni mint az *Apriori*.

6.2. Erősen korrelált adatok

A C20 és MUSHROOMS adathalmazokban a gyakori generátorok száma sokkal kevesebb, mint az összes gyakori minta száma. Emiatt – köszönhetően a mintaszámláló következtetésnek – a *Zart* algoritmusnak sokkal kevesebb minta support-ját kell megállapítania mint az *Apriori*-nak. Valamennyi esetben megfigyelhető, hogy a *Zart* és *Pascal*-algoritmusok futási ideje csaknem azonos, vagyis a *Zart* algoritmus extra jellemzői (ügymint zárt minták azonosítása és a zárt minták generátorainak a megkeresése) nem okoznak szinte semmiféle teljesítménycsökkenést. Az *Apriori* nagyon hatékony ritka adathalmazokon, erősen korrelált adatok esetében viszont a másik két algoritmus sokkal jobban teljesít.

7. Konklúzió és jövőbeli tervek

Ebben a cikkben a *Zart* nevű multifunkcionális mintakereső algoritmust mutattuk be, mely a *Pascal*-algoritmus kiterjesztése. A *Pascal*-tól eltérően a *Zart* azonosítani tudja a gyakori zárt mintákat, ill. megkeresi a zárt minták generátorait. Megmutattuk, hogy szükség van ezen extra tulajdonságokra a minimális nem redundáns asszociációs szabályok előállításához. A teszteredmények szerint a *Zart* a *Pascal*-al majdnem azonos válaszidőket produkál mind gyengén, mind pedig erősen korrelált adatokon. Tehát a többletkimenet nem megy a teljesítmény rovására.

Érdekes lenne megvizsgálni azt a kérdést, hogy vajon a *Zart*-ban bemutatott ötlet általánosítható-e. Vajon bármely gyakori mintákat kereső algoritmus kiterjeszthető ilyen formában, legyen az akár szélességi, akár mélységi kereső? Vajon lehetséges lenne úgy kiterjeszteni ezeket az algoritmusokat, hogy ne csak az összes érvényes szabályt, hanem közvetlenül a minimális nem redundáns szabályokat is elő lehessen velük állítani? A jövőben ezen kérdésekre szeretnénk választ találni.

Hivatkozások

- [1] BASTIDE, Y. AND TAOUIL, R. AND PASQUIER, N. AND STUMME, G. AND LAKHAL, L.: *Mining Minimal Non-Redundant Association Rules Using Frequent Closed Itemsets*. In: Proc. of the 1st Intl. Conf. on Computational Logic (CL '00). Volume 1861 of LNAI., Springer (2000) 972–986.
- [2] KRYSZKIEWICZ, M.: *Representative Association Rules*. In: Proc. of the 2nd Pacific- Asia Conf. on Research and Development in Knowledge Discovery and Data Mining (PAKDD '98). Volume 1394 of LNCS., Springer-Verlag (1998) 198–209.
- [3] GUIGUES, J. L., DUQUENNE, V.: *Familles minimales d'implications informatives résultant d'un tableau de données binaires*. Math. et Sci. Hum. **95** (1986) 5–18.
- [4] LUXENBURGER, M.: *Implications partielles dans un contexte*. Mathématiques, Informatique et Sciences Humaines **113** (1991) 35–55.
- [5] KRYSZKIEWICZ, M.: *Concise Representations of Association Rules*. In: Proc. of the ESF Exploratory Workshop on Pattern Detection and Discovery. (2002) 92–109.
- [6] PASQUIER, N., BASTIDE, Y., TAOUIL, R., LAKHAL, L.: *Efficient Mining of Association Rules Using Closed Itemset Lattices*. Inf. Syst. **24**(1) (1999) 25–46.
- [7] PASQUIER, N., BASTIDE, Y., TAOUIL, R., LAKHAL, L.: *Discovering Frequent Closed Itemsets for Association Rules*. LNCS **1540** (1999) 398–416.
- [8] ZAKI, M. J., HSIAO, C. J.: *CHARM: An Efficient Algorithm for Closed Itemset Mining*. In: Proc. of SDM '02. (2002) 33–43.
- [9] KRYSZKIEWICZ, M.: *Concise Representation of Frequent Patterns Based on Disjunction-Free Generators*. In: Proc. of ICDM '01, Washington, DC, USA, IEEE Computer Society (2001) 305–312.

- [10] BYKOWSKI, A., RIGOTTI, C.: *A Condensed Representation to Find Frequent Patterns*. In: Proc. of PODS '01, ACM Press (2001) 267–273.
- [11] KRYSZKIEWICZ, M., GAJEK, M.: *Why to Apply Generalized Disjunction-Free Generators Representation of Frequent Patterns?* In: Proc. of ISMIS 2002, Lyon, France, Springer-Verlag Berlin / Heidelberg (2002) 383–392.
- [12] PASQUIER, N.: *Mining Association Rules Using Formal Concept Analysis*. In: Proc. of ICCS '00, Shaker-Verlag (2000) 259–264.
- [13] BASTIDE, Y., TAOUIL, R., PASQUIER, N., STUMME, G., LAKHAL, L.: *Mining Frequent Patterns with Counting Inference*. SIGKDD Explor. Newsl. **2**(2) (2000) 66–75.
- [14] SZATHMÁRY, L.: *Symbolic Data Mining Methods with the Coron Platform*. PhD Thesis in Computer Science, University Henri Poincaré – Nancy 1, France (2006)
- [15] SZATHMÁRY, L., NAPOLI, A., KUZNETSOV, S. O.: *ZART: A Multifunctional Itemset Mining Algorithm*. In: Proc. of the 5th Intl. Conf. on Concept Lattices and Their Applications (CLA '07), Montpellier, France (2007) 26–37.
- [16] PASQUIER, N., BASTIDE, Y., TAOUIL, R., LAKHAL, L.: *Closed Set Based Discovery of Small Covers for Association Rules*. In: Proc. of BDA '99. (1999) 361–381.

(Beérkezett: 2009. augusztus 28.)

SZATHMÁRY LÁSZLÓ

Dépt. d'Informatique UQAM, C.P. 8888,
Succ. Centre-Ville, Montréal H3C 3P8, Canada
Debreceni Egyetem
Informatikai Kar
Debrecen H-4010 Pf. 12
Szathmary.L@gmail.com

BOGNÁR KATALIN

Debreceni Egyetem
Informatikai Kar
Debrecen H-4010 Pf. 12
bognar.katalin@inf.unideb.hu

ZART: A MULTIFUNCTIONAL ITEMSET MINING ALGORITHM

LÁSZLÓ SZATHMÁRY AND KATALIN BOGNÁR

In this paper, we present and detail a multifunctional itemset mining algorithm called *Zart*, which is based on the *Pascal* algorithm. *Zart* shows a number of additional features and performs the following, usually independent, tasks: identify frequent closed itemsets and associate generators to their closures. This makes *Zart* a complete algorithm for computing classes of itemsets

including generators and closed itemsets. These characteristics allow one to extract minimal non-redundant association rules, a useful and lossless representation of association rules. In addition, being based on the *Pascal* algorithm, *Zart* has a rather efficient behavior on weakly and strongly correlated data. Accordingly, *Zart* is at the heart of the *CORON* platform, which is a domain independent, multi-purposed data mining platform, incorporating a rich collection of data mining algorithms.

EGY ÚJ FELADAT: LÁDAFEDES SZÁLLÍTÁSSAL ÉS ENNEK MEGOLDÁSA ALGORITMUSOK EVOLÚCIÓJÁVAL

BENKŐ ATTILA, DÓSA GYÖRGY

A cikkben egy új feladatot definiálunk, amelyet „*Ládafedés szállítással*”-nak nevezünk. A feladat egyik jellegzetessége, hogy nem csak jó, hanem „*jó és gyors*” pakolást, vagy fedést keresünk. Néhány algoritmust adunk meg ilyen típusú (nagyon nehéz) feladatok megoldására, és egy új módszert is bemutatunk, amelyet „*Algoritmusok evolúciójá*”-nak nevezünk. Ezen a következőt értjük: Definiálunk egy algoritmuscsaládot, amely képes megoldani a feladatot, utána egy szomszédsági struktúrát ezen algoritmusok között, majd egy metaheurisztikát használunk (ebben a cikkben szimulált hűtést) a legmegfelelőbb (legjobb megoldást adó) algoritmus kiválasztására. Számítógépes tesztek segítségével demonstráljuk a módszer hatékonyságát.

1. Bevezető

Egy új problémával foglalkozunk, amit „*Ládafedés szállítással*”-nak nevezünk. Ezen típusú feladatok esetében tárgyakat pakolunk ládába, majd a lezárt ládákat (ahova már nem akarunk további tárgyakat pakolni) elszállítjuk. A feladatot a [8] cikkben definiáltuk, és megadtuk annak néhány lehetséges változatát. Itt most a lehetséges változatok közül csak egyet vizsgálunk. Ezt a későbbiekben pontosan definiáljuk, előljáróban csak annyit, hogy a fedett (és elszállított) ládákért pénzt kapunk, és ezt a profitot maximalizáljuk.

A tisztán ládapakolási feladat (BP) esetén adott a tárgyak mérete, a tárgyakat be kell pakolni minimális számú ládába, úgy, hogy a ládába pakolt tárgyak összmérete nem haladhatja meg a láda kapacitását, amit 1-nek szokásos választani. A ládafedési feladat (BC) esetén a ládát fedettnek tekintjük, ha a ládába pakolt tárgyak összmérete nem kisebb, mint a láda kapacitása (megint általában 1), és annyi ládát akarunk fedni, amennyit csak lehetséges. Közismert tény [3], hogy mindkét probléma megoldása NP-nehéz. Ez azt jelenti, hogy az optimális megoldás megtalálásához általában exponenciálisan sok lépés szükséges. Sok esetben nem áll rendelkezésünkre ennyi idő, viszont megelégszünk a feladatnak elég jó, közel-optimális megoldásával is.

Amikor a tárgyak egyesével jönnek, és amint megjelenik a következő tárgy, azonnal kénytelenek vagyunk azt valamely ládába pakolni, vagyis az online esetben, általában nem lehet olyan jó algoritmust konstruálni, mint az offline esetben, amikor előre ismerjük a tárgyakat.

Vegyük például a BC feladatot megoldó Duál Next Fit online algoritmust. Ez egyszerre csak egyetlen nyitott ládát használ, a következő tárgyat mindig az éppen nyitott ládába helyezi, amíg a ládába helyezett tárgyak összmérete kisebb mint a láda mérete. Amint megtelt a láda, lezárja a ládát, és egy új ládát nyit meg. Ennél az algoritmusnál a felhasznált ládák száma (a legrosszabb esetben) kétszer több is lehet, mint az offline optimális megoldás esetén.

Ebben a cikkben az új feladatnak csak az online esetével foglalkozunk, vagyis itt is feltesszük majd, hogy a tárgyak egyesével érkeznek. A tárgyakkal ládákat akarunk fedni (vagyis telepakolni), de a célfüggvényünk most nemcsak a fedés *jó-ságától*, de a fedési eljárás *gyorsaságától* is függ. Tehát, egyszerre „jó és gyors” fedést keresünk.

Számos módon lehet ilyen problémát definiálni: Egyik lehetőség lenne egy K méretű puffer használata, (lásd, ütemezési feladat esetén [1, 2]), ahol a puffer mérete, vagy a pufferben tárolt tárgyak száma növeli a célfüggvényt: bünteti a várakozást. Ebben a cikkben egy másik módot választunk: a célfüggvény „büntetése” a megnyitott ládák száma alapján történik, hiszen egyszerre több megnyitott láda kezelése több időt igényel.

Létezik néhány publikáció az „Ütemezés szállítással” témakörében, (lásd: [4], és az ebben hivatkozott cikkeket), de a ládapakolás vagy ládafedés szállítással kombinált változatát korábban még nem vizsgálták.

A cikk második fejezetében megadjuk a vizsgált feladat pontos definícióját, továbbá megvizsgáljuk néhány jellegzetes tulajdonságát.

A harmadik fejezetben bemutatunk néhány természetesen adódó algoritmust, a negyedik fejezetben definiáljuk az algoritmusok evolúcióját (EOA), és ennek hatékonyságát demonstráljuk. Néhány kiegészítő megjegyzéssel zárul a dolgozatunk.

2. A probléma definíciója

Számos lehetőség van, hogy definiáljuk ezt a problémát: „Ládafedés szállítással”, most ezekből csak egy lehetőséget választunk, ebben a cikkben csak ezzel foglalkozunk. A többi érdekes, ill. fontos változat vizsgálata további kutatás tárgya marad. A problémánk pontos megfogalmazása a következő:

Egyenként érkeznek a tárgyak, az i -dik tárgy mérete legyen $p_i > 0$. Adott továbbá egy $K > 0$ egész szám, egyszerre legfeljebb K számú láda lehet nyitva. A ládák kapacitása 1. Amint megérkezik egy tárgy, rögtön be kell hogy pakoljuk egy ládába. A tárgyat tehetjük egy már megnyitott ládába, vagy nyithatunk egy új ládát és a tárgyat ebbe az újonnan nyitott ládába is pakolhatjuk, de amint mondtuk, egyszerre legfeljebb csak K számú láda lehet nyitva. Amint megtelik egy láda (a láda megtelt, vagy más szóval fedett, ha a ládába pakolt tárgyak méretének összege legalább 1), azt azonnal elszállítjuk. A célfüggvény egy $G : \{1, \dots, K\} \rightarrow \mathbb{R}$ haszon-függvénnyel van megadva a következőképpen: ha éppen k a nyitott ládák száma (ahol $1 \leq k \leq K$) amikor egy láda megtelik, $G(k)$ profitot kapunk a láda

fedéséért. A G haszon-függvényről feltesszük, hogy pozitív, monoton nem növekvő függvény. Ebben a megközelítésben azt modellezzük, hogy több nyitott láda kezeléséhez több időre van szükségünk, hogy eldöntsük, hova pakoljuk az aktuális tárgyat. A cél a teljes haszon maximalizálása. (Ha végül marad fedetlen láda, azért nem jár pénz.) A tárgyak számát n -el jelöljük, de ezt előre nem ismerjük, csak amikor kiderül, hogy nem érkezik több tárgy.

Valós-életbeli alkalmazáshoz álljon itt a következő példa: egy kisebb konzervgyárban kézzel pakolják a gyümölcsöt dobozokba, a gyümölcs egy ablakon keresztül érkezik, egy rakodómunkás van az ablak túloldalán, aki ezt a feladatot végzi. Minden ládába legalább s összsúlyú gyümölcsöt kell pakolni, ennél többet szabad, de nyilván nem érdemes sokkal többet pakolni, mint amennyit muszáj.

Természetes azt feltételezni, hogy ez az ember nem tud túl sok nyitott ládát egyszerre kezelni, továbbá minél több nyitott ládával dolgozik, ez annál több időt igényel tőle, és a feladat egyre bonyolultabb lesz számára. Tehát a következő észrevételeket tehetjük:

Egyrészt, igyekeznie kell minél jobb fedést generálnia (vagyis mindegyik láda legyen telepakolva, de ne túlságosan: a ládába pakolt gyümölcsök összsúlya legyen legalább s , de azt ne nagyon haladja meg), ami nem könnyű feladat, minél kevesebb nyitott láda van, annál nehezebb. Másrészt mivel a gyorsaság is számít, érdeke, hogy egyszerre lehetőleg kevés számú láda legyen nyitva, vagyis kevés lehetőség közül kelljen választania.

Az előbbi két érdek egymásnak ellentmond, a dobozba pakoló személy célját ezért úgy szimulálhatjuk, hogy minden fedett ládáért pénzt kap (emiat érdekes sok ládát telepakolnia, vagyis jó fedést csinálnia), de másrészt ahogy a nyitott ládák száma növekszik, egyre kevesebb pénzt kap egy-egy fedett ládáért (vagyis érdemes kevés nyitott ládával dolgoznia).

Az online algoritmusok hatékonyságát rendszerint versenyképességi analízissel mérik. Ez azt jelenti, hogy egy A online algoritmus által kapott $C_A(I)$ célfüggvényértéket (ahol I -vel jelölik az inputot) összehasonlítják (elosztják) az offline optimum $OPT(I)$ értékével. Maximalizálási feladatok esetén (ahogy a mi esetünkben is) a $C_A(I)/OPT(I)$ hányados infimumát (ahol az infimumot tetszőleges I inputon vesszük) az A algoritmus versenyképességi hányadosának nevezzük. Mit értünk „offline” feladaton és mit „offline optimum”-on? Offline feladatok esetén mindig feltételezzük, hogy az inputra vonatkozó összes információ előre ismert. Ha az (ismert) tárgyakat az offline feladat esetén tetszőleges sorrendben szabad pakolni a ládába, akkor a G haszon-függvénynek semmi szerepe nincs.

(Ha a tárgyakat előre ismerjük, akkor – exponenciális időben – megadható az optimális fedés. Ezt a fedést egy lista szerint is el lehet készíteni – amelyik lista persze általában más lesz, mint az adott L lista – úgy, hogy egyszerre csak egy ládát kell nyitva tartani.)

Ezért úgy definiáljuk az offline modellt, hogy előre ismerjük a bemenetre vonatkozó összes információt, de a tárgyakat az adott L lista szerinti sorrendben kell a ládába pakolnunk (abban a sorrendben, ahogy azok valójában, online módon érkeznek).

Természetesen ilyen offline optimális megoldásnak léteznie kell. Bármelyik pillanatban, amikor egy új tárgy érkezik, véges sok lehetőség közül lehet választani (legfeljebb K számú ládába lehet a következő tárgyat pakolni). A mindent összevetve is véges számú lehetőség között léteznie kell olyan megoldásnak, ami a legjobb értékét adja a célfüggvénynek. Természetesen az offline-optimum nemcsak a tárgykészlettől függ, de az adott L listától is és a G haszon-függvénytől is.

Bármely véges L tárgy-lista és G haszon-függvény esetén legyen $C_A(L, G)$ egy A algoritmus által kapott megoldási értéke. Ezt hasonlítjuk össze az offline-optimális megoldással, amit $OPT(L, G)$ -vel jelölünk. Ekkor azt mondjuk, hogy az A algoritmus ρ -kompetitív, $(0 \leq \rho \leq 1)$ ha

$$\frac{C_A(L, G)}{OPT(L, G)} \geq \rho$$

teljesül bármely L, G esetén. Az előbbi ρ számok maximumát az A algoritmus versenyképességi hányadosának nevezzük. Másrészt, tegyük fel, hogy az L sorozat és G haszon-függvény esetén tetszőleges A online algoritmusra

$$\mu \geq \frac{C_A(L, G)}{OPT(L, G)}$$

teljesül valamilyen μ számmal, az ilyen μ számok legkisebbikét a feladat felső korlátjának nevezzük. Egy algoritmus optimális, ha a ρ versenyképességi-hányadosa meg-
egyezik a feladat μ felső korlátjával.

3. Néhány (természetesen adódó) algoritmus

Ebben a fejezetben először megadjuk néhány klasszikus algoritmus természetes adaptációját. Először tekintsük a Duál Next-Fit (röviden DNF) algoritmust. DNF mindig csak egy ládát tart nyitva, és amint a láda megtelik, elszállítjuk. Az algoritmikus leírás az alábbi:

DNF algoritmus

- A következő tárgyat mindig a nyitott ládába pakoljuk. Csak akkor szállítjuk a ládát, ha az fedett lesz, ekkor egy új ládát nyitunk a további tárgyak számára. Ha már nem jön több tárgy, az algoritmus megáll.

Ezen algoritmus alkalmazása esetén csak egy lehetőség van a soron következő tárgy pakolásához. Azonban ez az algoritmus mégis optimális tud lenni abban a speciális esetben, ha az elszállított ládák után kapott haszon nulla, ha több mint egy láda van egyszerre nyitva. Továbbá optimális abban az esetben is, ha a tárgyméretek majdnem egyformák, a következő lemma szerint:

3.1. LEMMA. *Tegyük fel, hogy $1/m \leq p_i < 1/(m-1)$ teljesül minden tárgyméretre, ahol $m \geq 2$ egész szám. Ekkor a DNF algoritmus optimális.*

Bizonyítás. Minden fedett láda pontosan m tárgyat tartalmaz. Emiatt a legjobb választás, hogy egyszerre csak egy láda van nyitva. \square

Természetesen általában nem feltételezhetjük a tárgyméretekre az előző lemma feltételének teljesülését. Ezért a DNF algoritmus alkalmazásakor veszteség keletkezik abban az értelemben, hogy számos láda „túl lesz pakolva”, amin azt értjük, hogy néhány láda úgy lesz lefedve, hogy a bele pakolt tárgyak összmérete jóval nagyobb lehet 1-nél, tehát jóval több, mint ami éppen elegendő lenne. Ha ezt a veszteséget meg akarjuk „menteni”, lehetővé kell tennünk, hogy egyszerre több láda legyen nyitható, ezáltal a ládaméret jobban közelíthetővé válik.

Tekintsük a következő klasszikus algoritmus, a *Harmonic*(K) algoritmus, vagy röviden $H(K)$ adaptációját, ahol $K \geq 1$. Az algoritmus elve az, hogy csak hasonló méretű tárgyak kerülhetnek egy ládába. Egyszerre legfeljebb csak K láda lehet nyitva. A tárgyakat a méreteik alapján osztályokba csoportosítjuk, minden láda egy osztályt reprezentál. Ha a következő tárgy mérete az $I_k = \left(\frac{1}{k+1}, \frac{1}{k}\right]$ intervallumba esik, a k -dik ládatípusba kerül, ahol: $k = 1, \dots, K-1$. A legkisebb tárgyak, vagyis az $I_K = \left(0, \frac{1}{K}\right]$ intervallumba esők pedig a K -dik típusú ládába kerülnek.

$H(K)$ algoritmus

- Helyezzük a következő tárgyat a k -dik típusú ládába, ha van ilyen nyitott láda, és a tárgy mérete az I_k intervallumba esik. Ha nincs ilyen nyitott láda, akkor nyitunk számára egy ilyen típusút. Amint egy láda megtelik (fedetté válik), elszállítjuk. Ha nincs további tárgy, az algoritmus megáll.

Megfigyelhető, hogy a $H(K)$ algoritmus jobban teljesít, mint a DNF, ha K „nem túl nagy” és a fedett ládák után járó profit-függvény „nem csökken túl gyorsan”. Megjegyezzük, hogy a $H(K)$ okosabbá tehető oly módon (vagyis nyerünk egy új algoritmust, amit Smart Harmonic(K)-nak, vagy röviden $SH(K)$ -nak neveztünk el, ami $H(K)$ „ügyes” változata), ha a következő szabályok szerint pakolja a tárgyakat:

$SH(K)$ algoritmus

1. Ha a következő elem le tudja fedni valamelyik ládát, akkor tegyük a tárgyat ezek közül a legkisebb telítettségű ládába, és szállítsuk el a ládát.
2. Bármely más esetben az $SH(K)$ algoritmus a $H(K)$ algoritmus szabálya szerint működik.

A következő táblázatban az előző algoritmusokat hasonlítottuk össze. Soronként a különböző algoritmusok által kapott megoldásokat közöljük, soronként más-más feladatosztályokra alkalmazva őket. A probléma-osztályok az alábbiak:

A ládaméretet minden esetben 100-nak, a tárgyak számát pedig 1000-nek választottuk. A tárgyméretek soronként a következő intervallumokból kerülnek ki: $[10; 40]$, $[15; 25]$, $[1; 100]$, és ismét $[10; 40]$, $[15; 25]$, $[1; 100]$. A profit-függvény az első három esetben: $G(k) = 11 - k$ (pl. $G(1) = 10$, $G(2) = 9$ és így tovább), ami egy gyorsan csökkenő profit-függvény. A következő három esetben pedig

$G(k) = 10, 1 - 0, 1k$ (pl. $G(1) = 10$, $G(2) = 9,9$ és így tovább), vagyis egy lassan csökkenő profit-függvényt választottunk.

Így nyerünk hat lényegesen különböző feladatosztályt. A DNF algoritmus által kapott megoldás értékét mindig 100%-nak tekintettük, és ehhez hasonlítottuk a többi algoritmus által kapott megoldásokat. Tíz futás átlagát írtuk a táblázat megfelelő rublikáiba. (A program C-nyelven íródott, a kódot kérésre szívesen elküldjük.)

| | DNF | H(2) | H(3) | SH(2) | SH(3) |
|---|-------------|--------|--------|--------|--------|
| 1 | 2080 = 100% | 98,6% | 97,4% | 99% | 97,5% |
| 2 | 1770 = 100% | 101,1% | 100% | 101,1% | 101,1% |
| 3 | 3460 = 100% | 93,6% | 83,1% | 99,2% | 100,6% |
| 4 | 2080 = 100% | 100% | 101,2% | 98,6% | 101,8% |
| 5 | 1770 = 100% | 100,6% | 100% | 100% | 101,1% |
| 6 | 3460 = 100% | 101% | 95,7% | 103,2% | 108,4% |

1. táblázat. Természetes algoritmusok

Emlékezzünk vissza, hogy az első három esetben a haszon-függvény erősen csökkenő. Az első feladatosztály esetén a többi algoritmus nem képes a DNF algoritmust „legyőzni”, de a második és a harmadik esetben már van olyan algoritmus, amelyik legyőzi a DNF-et. A többi esetben a haszon-függvény csak enyhén csökken, itt a DNF könnyen legyőzhető, de egy idő után $H(K)$, ill. $SH(K)$ hatékonysága nem javul tovább K növelésével, $SH(K)$ bizonyos esetekben tényleg „okosabb”, vagy „ügyesebb” mint a $H(K)$ algoritmus egyszerű változata (ugyanazzal a K -val). Figyeljük meg, hogy a hatodik feladatosztály esetében az $SH(K)$ lényegesen jobb, mint $H(K)$, ennek az lehet az oka, hogy a tárgyméreték eloszlása (1 és 100 között) nagyobb lehetőséget ad az algoritmusnak, hogy „okos” legyen.

3.1. Egy új, rugalmas algoritmus-család

Most egy új algoritmus-családot definiálunk, amely kellően rugalmas ahhoz, hogy az előbbi algoritmusok bármelyikével sikeresen felvegye a versenyt. Ezt különböző stratégiai paraméterek beállításával érjük el. Az algoritmust $Mask(\alpha, \beta, K)$ -val jelöljük, és az alábbiak szerint működik. K az egyszerre megnyitható ládák maximális számát jelenti. α és β K -dimenziós nemnegatív vektorok, ahol minden koordináta kisebb, mint 1. Az algoritmus egy elfogadás-elutasítás politikát folytat: a következő tárgyat elfogadja, és valamely ládába pakolja, ha a ládába pakolt tárgyak (az aktuális tárgy méretével) megnövelt összmérete az „elfogadó” tartományba kerül, és elutasítja a tárgynak a ládába történő pakolását, ha a megnövelt összméret az „elutasító” tartományba kerül. Az elutasító, illetve az elfogadó tartományokat az α és β paraméterek definiálják a következőképpen:

A k -dik láda elfogadó tartománya a következő: $[0; 1 - \alpha_k] \cup [1; 1 + \beta_k]$. Vagyis minden láda esetén két elfogadó és két elutasító tartomány van. Egyrészt, nem engedjük meg a tárgynak a ládába pakolását, ha az összméret ezáltal túlságosan nagy lenne, nagyon „túl lenne pakolva” a láda, vagyis, ha az összméret $(1 + \beta_k)$ -nél nagyobbá nőne. Másrészt azt sem engedjük meg, ha a megnövelt összméret már majdnem 1, de még kevesebb, mint 1 (vagyis a láda még nincs tele, de már majdnem tele van). Ilyenkor ugyanis, megfelelően nagy tárgyméretek esetén a következő tárgy, amelyik megtölti a ládát, egyben jelentősen „túlpakolná” azt. Ezek alapján a $\text{Mask}(\alpha, \beta, K)$ algoritmus formális leírása a következő:

***Mask*(α, β, K) algoritmus**

1. Ha a következő tárgy lefedi valamelyik ládát az elfogadó-tartományban, akkor pakoljuk a tárgyat abba a ládába, amelyek ezen ládák közül a legkisebb telítettségű. Szállítsuk el a ládát és menjünk az 5-ös pontra.
2. Ha a következő tárgy pakolható valamelyik ládába (az elfogadó-tartományban, de a láda meg nem lesz fedett), akkor pakoljuk azt egy ilyen ládába. Menjünk az 5-ös pontra.
3. Ha $k < K$, akkor nyissunk egy új ládát, az aktuális tárgy ebbe a ládába kerül, és menjünk az 5-ös pontra.
4. Ha $k = K$, akkor pakoljuk az aktuális tárgyat a legkisebb telítettségű ládába. Ha a láda fedett lesz, szállítsuk el és menjünk az 5-ös pontra.
5. Ha nincs több tárgy, az algoritmus megáll, különben menjünk az 1-es pontra.

Megjegyzés. A *Mask* algoritmusnak itt már rögtön az „ügyes” változatát definiáltuk, vagyis ha a következő tárggyal be tudunk fedni egy ládát, az elfogadási tartományon belül, ezek közül olyanba tesszük, amelyik a legkevésbé lesz túlpakolva.

A következő részben a *Mask* algoritmus stratégiai K, α, β paramétereinek „jó” megválasztásával foglalkozunk. Azt fogjuk látni, hogy lényegében minden feladatosztály esetén be lehet úgy állítani a paramétereket, hogy a megfelelő beállítás mellett *Mask* legalább olyan jó, vagy még hatékonyabb, mint az előző algoritmusok.

Ugyanazt a hat feladatosztályt vizsgáltuk, mint az előbb, a „max” oszlopában a korábbi hét algoritmus (DNF, H(2), H(3), H(4), SH(2), SH(3), SH(4)) által kapott legjobb megoldás értéke szerepel (%-os formában), ezzel versenyeztetünk különböző paraméterű *Mask* algoritmusokat. A paraméterek a következők:

*Mask*₁: $K = 2$, $\alpha = (15, 20)$, és $\beta = (10, 30)$.

*Mask*₂: $K = 1$, $\alpha = (10)$, és $\beta = (20)$.

*Mask*₃: $K = 2$, $\alpha = (15, 15)$, és $\beta = (30, 30)$.

*Mask*₄: $K = 3$, $\alpha = (20, 20, 20)$, és $\beta = (30, 30, 30)$.

*Mask*₅: $K = 3$, $\alpha = (10, 10, 10)$, és $\beta = (25, 25, 25)$.

*Mask*₆: $K = 3$, $\alpha = (10, 20, 30)$, és $\beta = (40, 50, 60)$.

| | max | $Mask_1$ | $Mask_2$ | $Mask_3$ | $Mask_4$ | $Mask_5$ | $Mask_6$ |
|---|---------|----------|----------|----------|----------|----------|----------|
| 1 | 100 % | 100, 2% | 100, 1% | 96, 2% | 91% | 93, 7% | 88, 9% |
| 2 | 101, 1% | 91, 5% | 101, 1% | 94, 6% | 86, 2% | 88, 8% | 88, 3% |
| 3 | 100, 6% | 100% | 101, 5% | 106% | 100% | 103, 9% | 100, 6% |
| 4 | 104, 3% | 105, 9% | 100, 5% | 104% | 106, 6% | 102% | 102, 8% |
| 5 | 101, 1% | 100, 8% | 101, 1% | 103, 1% | 102, 6% | 103, 5% | 102, 3% |
| 6 | 108, 4% | 105% | 101, 5% | 108, 3% | 113, 1% | 110% | 114% |

2. táblázat. Mask algoritmusok különböző stratégiai paraméterekkel

Természetesen nem mindegyik Mask „jó” minden esetben. Ez nem is lenne célunk. Azonban megfigyelhetjük a következőt: mindegyik feladatosztály esetén **van olyan Mask algoritmus** (van olyan paraméter-beállítás), amelyik versenyképes a korábbi legjobb algoritmussal, sőt legyőzi azt. Azt kell még „kitalálnunk”, hogy hogyan tudjuk megtalálni, egy adott feladatosztály esetén a paraméterek megfelelő beállítását. Ezzel a kérdéssel foglalkozunk a következő fejezetben.

4. Algoritmusok evolúciója - EOA

Megállapítottuk tehát, hogy a Mask algoritmus paramétereinek helyes megválasztásával képesek vagyunk a feladatot jól megoldani és Mask felülmúlja a többi, korábban tárgyalt algoritmust. Az egyetlen problémát a paraméterek „helyes” megválasztása jelenti. Ebben a fejezetben egy új módszert adunk a ládáfedési feladatunk megoldására, a módszert „Algoritmusok evolúciójának” (EOA) nevezzük. E módszer más (nehéz) online feladat megoldására is alkalmazható. Módszerünk nem azonos azzal, amit „Evolúciós algoritmus”-nak (EA) neveznek. Mi a fő különbség a kettő között? Az evolúciós algoritmusok (EA) (vagy más lokális kereső módszerek), valamilyen **offline** feladat megoldására szolgálnak. Legyen S a feladat megengedett megoldásainak halmaza, ezek között egy szomszédsági struktúrát definiálunk (a megoldásokat általában $0 - 1$, vagy valós vektorokkal reprezentáljuk). Az EA először meghatározza a feladat egy x_0 megoldását, majd kiindulva ebből az x_0 -ból, ennek a környezetéből választ egy (másik) x_1 megoldást, ha bizonyos kritériumok teljesülnek, akkor x_0 -át x_1 -re cseréli és ezt a lépést iterálja.

Most másról van szó, először is, feladatunk nem offline, hanem **online**. (Vagyis bizonyos értelemben az EOA az EA online megfelelője.) Nem az offline feladat megengedett megoldásai között, hanem az online feladat megoldó algoritmusai között hozunk létre szomszédsági struktúrát. Megengedett megoldások helyett algoritmusokon „lépegetünk”. A szerzők legjobb tudomása szerint ilyen módszert korábban még nem adtak meg és nem vizsgáltak. Reméljük, hogy módszerünk (meghatározzunk egy kellően rugalmas algoritmus-családot valamely online feladat megold-

dására, az algoritmusok között egy szomszédsági struktúrát definiálunk, eztán egy lokális kereső módszer segítségével kiválasztjuk az algoritmusok „legjobbikát”) alkalmazható lesz a jövőben más (nehéz) feladatok megoldására is. A főbb különbségeket EA és EOA között az alábbi táblázat szemlélteti.

| | EA | EOA |
|-------------------|----------------------|--------------------|
| a feladat jellege | offline | online |
| „amin lépegetünk” | megengedett megoldás | megoldó algoritmus |

3. táblázat. EA és EOA összehasonlítása

A szomszédsági struktúra természetes módon határozható meg a különböző Mask algoritmusok között: Egy rögzített $Mask(K, \alpha, \beta)$ algoritmusnak módosítjuk valamelyik paraméterét: K -t 1-gyel növeljük vagy csökkentjük, illetve az α vagy β vektorok valamelyik komponensét csökkentjük, vagy növeljük egy előre rögzített kicsi Δ pozitív konstanssal (úgy, hogy a megváltozott érték ismét pozitív legyen és kisebb, mint 1). A megfelelő szomszéd választásához szimulált hűtést alkalmaztunk, amely módszert az alábbiakban röviden összefoglaljuk.

A szimulált hűtés (simulated annealing) egy népszerű, sok feladat megoldására alkalmazott, hatékony metaheurisztika, közeli rokonságban áll a Tabu-kereséssel és a genetikus algoritmusokkal. Szimulált hűtés esetén adott egy (nehéz) kombinatorikus optimalizálási offline feladat, amelynek ismerjük egy x megengedett megoldását, esetleg egy ilyen megoldást valamilyen heuristikával állítunk elő. Feltételezzük, hogy minden megengedett megoldásnak könnyen elő tudunk állítani véletlenszerűen valamilyen y „szomszédját”, vagyis egy olyan másik megengedett megoldást, amely az előbbinek egy „kicsi” változtatásával adódik, vagyis kicsi környezetéből való. Szimulált hűtés esetén, ezután, a két megoldás közül az egyiket választjuk majd, megfelelő szabály szerint. Így, iterációnként egy-egy megengedett megoldásunk van, és ezekkel igyekszünk valamilyen optimális megoldást megközelíteni.

A szimulált hűtés azt szimulálja, amikor valamilyen anyagot felmelegítünk, majd lassú hűtés folyamán a részecskék valamilyen alaphelyzetbe kerülnek. A kulcs most az, hogy hogyan választunk az előbbi két megengedett megoldás, x és y közül. Egyrészt, ha y jobb megoldás, mint x , vagyis maximalizálandó célfüggvény esetén nagyobb célfüggvény-érték tartozik hozzá, akkor mindenképpen a jobb megoldást, y -t választjuk. Azonban y -t akkor is elfogadjuk, ha csak egy „kicsivel” rosszabb megoldás mint x , de egyre csökkenő valószínűséggel, és egyre kisebb célfüggvényromlási engedünk meg. A pontos képletek helyett, az érdeklődő olvasó figyelmébe inkább az [5] összefoglaló cikket ajánljuk. Szimulált hűtés esetén is sok múlik bizonyos stratégiai paraméterek megfelelő beállításán, amelyek az algoritmus hatékonyságát jelentősen befolyásolják.

Alább ismét közlünk egy összehasonlító tesztet, ahol azt vizsgáltuk, hogy vajon az EOA algoritmus által jobb eredményeket kapunk-e, mint a korábbi algoritmu-

sokkal (vagyis tényleg sikerül-e a Mask algoritmus stratégiai paramétereit helyesen beállítani, „kézi vezérlés”, vagyis próbálkozás nélkül).

A problémaosztályok ugyanazok voltak, mint korábban is. Lássuk a kapott eredményeket. (Egy „sima” lokális keresést (LS) is lefuttatunk minden esetben, amelynél csak akkor lépünk át a szomszéd algoritmusra, ha az az előzőnél jobb eredményt produkál.)

| | max | LS | EOA |
|---|------------|-----------|------------|
| 1 | 100, 2% | 100, 6% | 101% |
| 2 | 101, 1% | 101, 1% | 101, 7% |
| 3 | 106% | 106, 2% | 106, 9% |
| 4 | 106, 6% | 108, 7% | 109, 6% |
| 5 | 103, 5% | 103, 6% | 104, 7% |
| 6 | 114% | 116, 9% | 116, 9% |

4. táblázat. Algoritmusok evolúciója

Amint a táblázatból látható, az EOA módszerünk az adott feladat esetén tényleg működik, vagyis alkalmas arra, hogy a feladatot hatékonyan megoldó algoritmust konstruáljon. A más algoritmusok által kapott megoldásokat helyenként lényegesen sikerült javítani. A lokális keresés is sok esetben hatékonynak bizonyul, de ahogy várható volt, ennél EOA (szimulált hűtést alkalmazva) még hatékonyabb.

Természetesen módszerünk csak akkor képes jól működni, ha a tárgyak eloszlása az időben nem változik. Az algoritmus-család (Mask) tagjait pedig nem a teljes tárgyhalmazra, hanem annak egy-egy $m < n$ darabból álló ugyanakkora szeletére alkalmazzuk.

5. Következtetések

Cikkünkben definiáltunk egy új feladatot, aminek a „Ládafedés szállítással” nevet adtuk. Megadtunk néhány természetesen adódó algoritmust a feladat megoldására, aztán egy sokkal hatékonyabb módszert, amit „Algoritmusok evolúciója”-nak, vagy röviden EOA-nak nevezünk. Ez különbözik az evolúciós algoritmusoktól abban az értelemben, hogy itt nem egy offline feladat megengedett megoldásain lépegetünk, hanem egy online feladat megoldó algoritmusain (a stratégiai paraméterek változtatásával).

A lépegetéshez szimulált hűtést alkalmaztunk, hogy megtaláljuk a megfelelő paramétereket. Természetesen ehelyett tabu-keresés vagy genetikus algoritmus alkalmazását is érdemes lenne kipróbálni, ez további kutatás tárgya.

Mivel a feladat nagyon nehéz, a szokásos versenyképességi analízis elvégzése szinte lehetetlen, ezért számítógépes tesztekkel igazoltuk algoritmusunk hatékony-

ságát. Az EOA során egy kellően rugalmas algoritmus-család (esetünkben a Mask) közül kiválasztjuk azt, amelyik a „legalkalmasabb” a család tagjai közül a feladat megoldására egy adott feladatosztályon belül.

Kíváncsian várjuk, hogy a valós életbeli problémák mely területein, milyen egyéb „nehéz” problémák megoldására lesz sikeresen alkalmazható az újonnan bemutatott módszer.

Hivatkozások

- [1] G. DÓSA, L. EPSTEIN: *Online scheduling with a buffer on related machines*, J. of Comb. Optim., DOI 10.1007/s10878-008-9200-y
- [2] M. ENGLERT, D. OZMEN, M. WESTERMANN: *The power of reordering for online minimum makespan scheduling*, In: Proc 48th symp foundations of computer science (FOCS), (2008) 603–612.
- [3] M. R. GAREY, D. S. JOHNSON: *Computers and intractability*, W. H. Freeman and Company, New York (1979)
- [4] W. ZHONG, G. DOSA, Z. TAN: *On the machine scheduling problem with job delivery coordination*, European Journal of Op. Res., **182**, (2007) 1057–1072.
- [5] PETER J. M. VAN LAARHOVEN, EMILE H. L. AARTS: *Simulated annealing: theory and applications*, Mathematics and its applications, Kluwer Academic Publishers (1987)
- [6] ED: KAISA MIETTINEN, MARKO M. MAKELA, PEKKA NEITTANMAKI, JACQUES PÉRIAUX: *Evolutionary algorithms in engineering and computer science*, John Wiley and Sons Ltd (1999)
- [7] T. NEMETH, C. IMREH: *Parameter Learning Algorithms in Online Scheduling*, Conference of PhD Students in Computer Science, inf.u-szeged.hu
- [8] A. BENKŐ, G. DÓSA, ZS. TUZA: *Bin packing/covering with delivery*, manuscript (2009)

(Beérkezett: 2009. november 25.)

BENKŐ ATTILA

Pannon Egyetem

Matematika Tanszék

benko.attila@almos.vein.hu

DÓSA GYÖRGY

Pannon Egyetem

Matematika Tanszék

dosagy@almos.vein.hu

BIN COVERING WITH DELIVERY
AND SOLVING IT WITH EVOLUTION OF ALGORITHMS

ATTILA BENKŐ, GYÖRGY DÓSA

We deal with a new problem called Bin Covering with Delivery. Mainly we mean under this expression that we look for not only a good, but a *“good and fast”* covering. There are several ways to treat such a problem, but we investigate here only one of them, an online problem, which has real-life application, as well. We apply the appropriate version of some classical bin covering algorithms for the problem, and also propose a new method that we call “Evolution of Algorithms”, to solve this (algorithmically very hard) problem. In case of such methods a neighborhood structure is defined among the algorithms in a flexible algorithm-family, and using a metaheuristic (simulated annealing now) in some sense the best algorithm is chosen from the set of the algorithms, to solve the problem. We show the efficiency of the proposed method by several computer tests.

A LEGHOSSZABB SZÉRIÁK VIZSGÁLATA ^{1 2}

FAZEKAS ISTVÁN, KARÁCSONY ZSOLT³, LIBOR JÓZSEFNÉ

1. Bevezetés

Számos mű foglalkozott már a címben megjelölt témával, bár a szóhasználatban léteznek eltérések. Vannak szerzők, akik leghosszabb futamnak, vagy siker sorozatnak, vagy egyszerűen leghosszabb sorozatnak nevezik egy adott kísérletsorozatban az egymást követő azonos jelek leghosszabb szériáját. Például az érmedobás kísérletben az egymás után következő – vagyis írással meg nem szakított – fejdobások számának a maximumát leghosszabb fejszériának fogjuk nevezni. Ebben a munkában ismert rekurziós és aszimptotikus tételeket, valamint a szimuláció szolgáltatja eredményeket fogjuk összehasonlítani. Így Erdős-Révész [4] és Földes [5] aszimptotikus eredményeit fogjuk összevetni Schilling [15], Bloom [3] és Kopocinski [8] általunk esetenként kiegészített és bizonyított rekurzív formuláival, valamint a szimulációs eredményeinkkel. Megjegyezzük, hogy Binswanger-Embrechts [2] az aszimptotikus tételeket vetette össze szimulációs eredményekkel. Ebben a cikkben a rekurziós eljárásokat hangsúlyozzuk, ezért ezeket részletesen bizonyítjuk. A rekurziós eljárások adják a pontos eredményt. Az aszimptotikus eredmények csak hosszú dobássorozat esetén adnak jó közelítést. A szimulációs eredmények pedig véletlenszerűek, és a dobássorozat nagyon sokszori számítógépes legenerálása esetén közelítik a pontos értéket. Tanulmányunk kiterjed a szabályos és szabálytalan érme esetére is, valamint mindkétféle érménél vizsgáljuk a leghosszabb fejszéria és a leghosszabb bármilyen (tisztá fej vagy tisztá írás) széria hosszát is. A szimulációkat a MATLAB programmal végeztük 20000 ismétlést alkalmazva, rövid ($n = 50$) és hosszú ($n = 1000$) sorozatok esetén. Munkánk során vizsgáltuk a visszatevés nélküli húzásokat is, melyet például a kártyalap-húzás kísérlettel tudunk szemléltetni. A kérdés itt is ugyanaz, hogyan alakul a leghosszabb azonos jelsorozat hossza (például a leghosszabb tisztá piros lapszéria hossza a francia kártya csomagból történt húzások során). Írásunkban megemlítünk néhány matematikatörténeti, didaktikai érdekességet is, melyek segítenek az egyetemi, főiskolai hallgatók érdeklődését felkelteni a téma iránt.

¹ 2000 *Mathematics Subject Classification*: 97K50, 60C05, 60F05.

² *Key words and phrases*: széria, rekurzió, szimuláció

³ A bemutatott kutató munka a TÁMOP-4.2.1.B-10/2/KONV-2010-0001 jelű projekt részeként az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

2. Független kísérletsorozat (visszatevéses mintavétel)

Bevezetésül tekintsük Varga Tamás egy érdekes kísérletét, melyet Révész Pál 1978-ban ismertetett Helsinkiben egy nemzetközi matematikai konferencián [12], majd Schilling [15] cikkének bevezetéseként is szolgált. Varga a tanulócsoporthat két részre osztotta, majd az egyik csoportnak azt adta feladatuk, hogy mindenki dobjon fel 200-szor egy pénzérmét, és jegyezze le a kapott fej-írás eredményeket. A csoport másik részének pedig a kísérletet csak gondolatban kellett elvégezni, és a gondolati eredményeket lejegyezni. Vagyis nekik olyan 200 elemű fej-írás sorozatot kellett írniuk, amilyen szerintük egy 200 elemű dobássorozat. A munka végeztével összekeverték a lejegyzett eredményeket tartalmazó lapokat, majd átadták Vargának, aki majdnem 100%-os biztonsággal megmondta, hogy az adott lap valós eredményt tükröz-e, vagy kitaláltat. Hiszen míg a valós sorozatokban nem volt ritka a 7 (esetleg 8) egymást követő fej – Rényi Alfréd jól ismert $\log_2 200$ -as eredményével összhangban –, addig a képzelt sorozatokban maximum 5 egyforma elemet mertek a tanulók egymás után leírni.

Amikor volt szerencsénk Révész professzor úrral találkozni és beszélgetni erről, elmondta a kísérlet továbbvitelét is. Ő, miután ismertette a hallgatókkal a Varga-féle kísérletet, és az eredményt is megbeszélték, újra elvégeztette az eredeti kísérletet. Vagyis a hallgatók fele újra valós kísérletet végzett az érme 200-szori feldobásával, míg a csoport másik fele a gondolati eredményeit írta le. Az összegyűjtött papírlapokat újra sikerült majdnem teljes pontossággal szétválogatnia Révésznek. A magyarázat nagyon egyszerű. A hallgatók többsége csak az egyik-féle, például a leghosszabb fejszériára koncentrált, de már nem figyelt a leghosszabb írásszériára. Felvetődik tehát az alábbi két kérdés.

- i. Egy n hosszúságú sorozat esetén hogyan alakul a *leghosszabb fejszéria hossza*?
- ii. Egy n hosszúságú sorozat esetén mekkora a *leghosszabb bármilyen széria (akár fej, akár írás) hossza*?

2.1. Szabályos pénzérme esete

2.1.1. Leghosszabb fejszéria vizsgálata

Schilling [15] cikke nyomán vizsgáljuk a következőt. Dobjunk fel egy szabályos pénzérmét n -szer. Fejszériának nevezzük az egymást követő (tehát írással meg nem szakított) fejek sorozatát. Jelölje R_n a leghosszabb fejszéria nagyságát. Az eloszlásfüggvényünk: $F_n(x) = P(R_n \leq x)$. $F_n(x)$ -et elegendő nemnegatív egész x -ekre megadni (hiszen $F_n(x) = 0$, ha $x < 0$; $F_n(x) = F_n([x])$, ha $x \geq 0$, ahol $[x]$ jelöli x egészrészét). Legyen $A_n(x)$ azon n hosszúságú sorozatok száma, amelyekben a leghosszabb fejszéria nem haladja meg x -et. Szabályos érme esetén egy n elemű sorozatot vizsgálva kapjuk tehát:

$$F_n(x) = P(R_n \leq x) = \frac{A_n(x)}{2^n}.$$

De hogyan tudnánk meghatározni az $A_n(x)$ értékét?

Vegyük először azt az esetet, amikor a leghosszabb fejszéria legfeljebb 3 elemű ($x = 3$). Ha $n \leq 3$, akkor $A_n(3) = 2^n$, hiszen minden lehetséges eset megfelel annak a kritériumnak, hogy az egymás utáni fejek száma maximum 3. Ha viszont az $n > 3$, akkor a számunkra kedvező sorozatok kezdődhetnek a következőképpen (ahol F jelöli a fej, I pedig az írás dobást): I, FI, FFI, FFFI, és utánuk csak olyan jelsorozat van, amelyben nincs háromnál hosszabb fejszéria. Kapjuk tehát a következőt:

$$A_n(3) = A_{n-1}(3) + A_{n-2}(3) + A_{n-3}(3) + A_{n-4}(3), \quad \text{ha } n > 3.$$

Ugyanezzel a módszerrel adódik az általános rekurziós képlet.

2.1. ÁLLÍTÁS. (Schilling [15], 198. o.)

$$A_n(x) = \begin{cases} \sum_{j=0}^x A_{n-1-j}(x), & \text{ha } n > x, \\ 2^n, & \text{ha } 0 \leq n \leq x. \end{cases}$$

Megjegyzés. Ha megnézzük $A_n(1)$ értékeit, vagyis azon n elemű sorozatok számát, melyekben legfeljebb 1 hosszúságú fejszéria van, éppen a Fibonacci-sorozat (azaz $a_0 = 0$, $a_1 = 1$, és $a_n = a_{n-1} + a_{n-2}$, ha $n \geq 2$) 2-vel eltoló elemeit kapjuk.

| n | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | ... |
|----------|---|---|---|---|---|----|----|----|----|-----|
| $A_n(1)$ | 1 | 2 | 3 | 5 | 8 | 13 | 21 | 34 | 55 | ... |

A k -rendű Fibonacci-számok segítségével kifejezhető $A_n(k)$, sőt a k -rendű Fibonacci-polinomok felhasználásával a szabálytalan pénzérme esete is kezelhető, lásd [10].

R_n aszimptotikus viselkedését Földes Antónia [5] alábbi tétele alapján írhatjuk le.

2.1. TÉTEL. (Földes [5].) Valamennyi egész k esetén

$$P\left(R_n - \left\lfloor \frac{\log n}{\log 2} \right\rfloor < k\right) = \exp\left(-2^{-(k+1-\{\frac{\log n}{\log 2}\})}\right) + o(1), \quad (1)$$

ahol $[a]$ jelöli az egészrészét a -nak és $\{a\} = a - [a]$, a törtrésze.

Itt \log a természetes alapú logaritmust jelöli.

2.1.2. Leghosszabb bármilyen széria vizsgálata

Ismét alapul vesszük Schilling [15] cikkét. Dobjunk fel egy szabályos pénzérmét n -szer, és jelölje R'_n a leghosszabb széria (akár a tiszta fej, akár a tiszta írás) nagyságát. Legyen $B_n(x)$ azon n hosszúságú sorozatok száma, amelyekben a leghosszabb (tetszőleges) széria nem haladja meg x -et. Szabályos érme esetén egy n elemű sorozatot vizsgálva, kapjuk az eloszlásfüggvényt:

$$F'_n(x) = P(R'_n \leq x) = \frac{B_n(x)}{2^n}.$$

Most Schilling [15], 199. oldalon leírt ötletét használjuk. A fej-írás sorozat minden eleme alatt jelölje A azt, hogy az utána következő vele azonos, illetve K azt, hogy különböző. Például:

$$\begin{array}{cccccccccccc} F & F & F & I & F & I & F & I & I & I & I & F & F \\ A & A & K & K & K & K & K & A & A & A & K & A \end{array}$$

Az alsó A, K elemekből álló sorozatban a leghosszabb tiszta A sorozat akkor és csak akkor $k - 1$ hosszúságú, ha fölötte a leghosszabb tiszta széria (fej vagy írás) k hosszúságú. Ha a felső sorozat n hosszú, és ebben a leghosszabb széria k elemű, akkor az alsó sorozat $n - 1$ hosszú, és a leghosszabb A széria $k - 1$ elemű. Vagyis

$$B_n(x) = 2A_{n-1}(x - 1)$$

(hiszen minden alsó sorozat pontosan 2 felső sorozathoz tartozhat). Ennek felhasználásával kapjuk:

$$\begin{aligned} F'_n(x) = P(R'_n \leq x) &= \frac{B_n(x)}{2^n} = \frac{2A_{n-1}(x - 1)}{2^n} = \frac{A_{n-1}(x - 1)}{2^{n-1}} = \\ &= P(R_{n-1} \leq x - 1) = F_{n-1}(x - 1). \end{aligned}$$

Beláttuk tehát, hogy

$$F'_n(x) = F_{n-1}(x - 1), \quad (2)$$

vagyis visszaveztük esetünket a tiszta fejszéria esetére.

Most vizsgáljuk azt, hogy a leghosszabb széria pontosan k hosszúságú. Jelölések:

$b_n(k)$: n dobásból hányszor lesz a leghosszabb széria (akár fej, akár írás) *pontosan* k hosszúságú,

$a_n(k)$: n dobásból hányszor lesz a leghosszabb fejszéria *pontosan* k hosszúságú.

A $b_n(k)$ -ra Szászné Simon Judit is ad [18] doktori értekezésében rekurzív képletet, melyet mi kétféleképpen bizonyítunk. Először a (7) képletből vezetjük le, majd teljes eseményrendszerre bontással kapjuk meg az adott képletet. A (7) formulát a következő szakaszban igazoljuk csak, de természetesen ehhez nem fogunk támaszkodni a jelen szakaszra.

2.2. ÁLLÍTÁS. (Szászné [18].) Minden $n = 1, 2, \dots$ esetén $b_n(1) = b_n(n) = 2$, $b_n(r) = 0$, ha $r > n$ vagy $r \leq 0$

$$b_n(r) = \sum_{h=1}^{r-1} b_{n-h}(r) + \sum_{i=1}^r b_{n-r}(i), \quad \text{ha } 1 < r < n. \quad (3)$$

Első bizonyítás. Tekintsük a (7) összefüggésen a $p = q = \frac{1}{2}$ esetet. Mivel az összes elemi esemény száma 2^n , ezért a $p(n, k)$ kifejezés 2^n -nel való beszorzásával adódik a számunkra kedvező esetek száma, vagyis $a_n(k)$. Azaz (7)-ből

$$\underbrace{2^n p(n, k)}_{a_n(k)} = \sum_{j=0}^{k-1} \underbrace{2^{n-j-1} p(n-j-1, k)}_{a_{n-j-1}(k)} + \underbrace{2^{n-k-1} F_{n-k-1}(k)}_{A_{n-k-1}(k)}. \quad (4)$$

Alkalmazzuk újra Schilling ötletét, vagyis: a fej-írás sorozat minden eleme alatt jelölje A azt, hogy az utána következő vele azonos, és K pedig azt, hogy különböző. Ahogyan adódott $B_n(k) = 2A_{n-1}(k-1)$, ugyanúgy adódik $b_n(k) = 2a_{n-1}(k-1)$. Ennek felhasználásával (4) képletünk a következőképpen alakul:

$$\frac{b_{n+1}(k+1)}{2} = \sum_{j=0}^{k-1} \frac{b_{n-j}(k+1)}{2} + \frac{B_{n-k}(k+1)}{2}.$$

Végezzük el a 2-vel való beszorzást és alkalmazzuk a következő átindexelést:
 $k+1 \rightarrow r, \quad n+1 \rightarrow n, \quad j+1 \rightarrow h$. Kapjuk a bizonyítandó formulát.

Második bizonyítás. Nézzük meg, hogy mit is fejez ki a (3) jobb oldalán lévő két összeg. Bontsuk fel az eseményterünket aszerint, hogy milyen hosszú széria van elől.

Ha a sorozatunk h ($h = 1, 2, \dots, r-1$) hosszúságú szériával kezdődik, akkor a maradék $n-h$ dobásból kell a leghosszabb szériának r hosszúságúnak lenni. Ha a kezdő h széria fej, akkor a $(h+1)$ -edik elem írás kell hogy legyen. Ha a kezdő h széria írás, akkor a $(h+1)$ -edik elem fej kell hogy legyen.

$$\begin{array}{cc} \underbrace{F \dots F}_{h \text{ db fej}} & \underbrace{I \dots F \dots F \dots}_{n-h \text{ elem között}} \\ (1 \leq h < r) & r \text{ hosszú széria} \end{array}$$

vagy:

$$\begin{array}{cc} \underbrace{I \dots I}_{h \text{ db írás}} & \underbrace{F \dots I \dots I \dots}_{n-h \text{ elem között}} \\ (1 \leq h < r) & r \text{ hosszú széria} \end{array}$$

De ennek a kettőnek a száma megegyezik, és az összegük éppen $b_{n-h}(r)$. Ha r hosszúságú szériával kezdődik a sorozat, akkor a maradék $n-r$ elemből $i = 1, 2, \dots, r$ hosszú széria lehet. Ha az első r elem fej, akkor az $(r+1)$ -edik írás kell, hogy legyen, míg ha az első r elem írás, akkor az $(r+1)$ -edik fej kell hogy legyen.

$$\begin{array}{cc} \underbrace{F \dots F}_{r \text{ db fej}} & \underbrace{I \dots F \dots F \dots}_{n-r \text{ elem között}} \\ & \text{legfeljebb } r \text{ hosszú széria} \end{array}$$

vagy:

$$\underbrace{\text{I} \dots \text{I}}_{r \text{ db írás}} \quad \underbrace{\text{F} \dots \text{I} \dots \text{I} \dots}_{n-r \text{ elem között}}$$

legfeljebb r hosszú széria

Ezen két részeset száma egyenlő, összegük pedig $\sum_{i=1}^r b_{n-r}(i)$. \square

A $b_n(r)$ -rel tehát megadtuk azon sorozatok számát, melyekben az n dobás során a leghosszabb bármilyen széria (akár fej, akár írás) pontosan r hosszúságú.

R'_n aszimptotikus viselkedésének vizsgálatához Földes [5] már idézett eredményét használjuk. (1) és (2) alapján kapjuk az alábbi.

2.2. TÉTEL. Valamennyi egész k esetén

$$P\left(R'_n - \left\lfloor \frac{\log(n-1)}{\log 2} \right\rfloor < k\right) = \exp\left(-2^{-(k - \{\frac{\log(n-1)}{\log 2}\})}\right) + o(1),$$

ahol - mint korábban - $[a]$ az a egészrészét, $\{a\}$ pedig az a törtrészét jelöli.

2.1.3. Numerikus eredmények

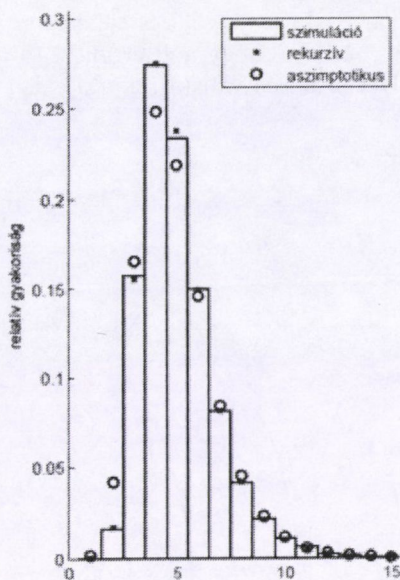
Vizsgálatunkhoz a MATLAB programot használtuk 20.000 ismétlésszámmal. Az alkalmazott számítógép paraméterei pedig a következők: INTEL Core Quad Q9550 processzor, 4Gb, DDR3 memória.

A következő ábrákon „x” jelöli a rekurzióval kapott eredményeket, „o” az aszimptotikus tételek eredményeit, az oszlopdiagram pedig a szimulációval kapott eredményeket mutatja. Az első grafikonpár a rövid ($n = 50$) sorozat eredményeit mutatja a leghosszabb fej (bal oldali) és a leghosszabb bármilyen (jobb oldali) széria esetén, majd a további grafikonpár ugyanezt a két esetet mutatja hosszú ($n = 1000$) dobáshosszra vonatkozóan.

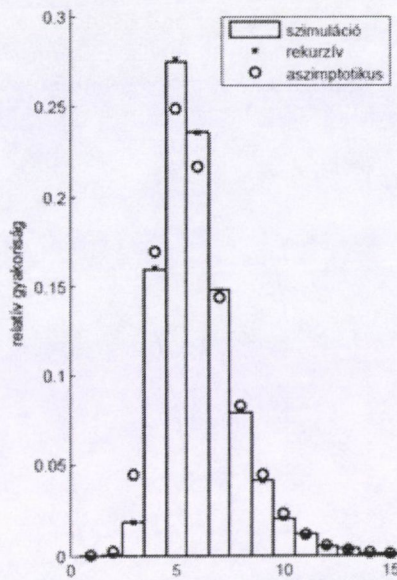
Mindkét esetre (leghosszabb fej, illetve leghosszabb bármilyen széria vizsgálatára) elmondható, hogy kis n esetén a szimulációs eredmények vannak közelebb a rekurzív eredményekhez, az aszimptotikus tételek n növelésével adják a rekurzióhoz közeli eredményeket. $n \geq 3000$ esetén a rekurziós, a szimulációs és az aszimptotikus értékek gyakorlatilag egybeesnek. Míg kis n esetén a rekurziós algoritmus gyors, n növelésével rohamosan lassul a számítási eljárás. A futási időket tekintve csak példaként néhányat megemlítve:

| n | ism | futási idő |
|--------|--------|---------------|
| 10.000 | 20.000 | 31.795.056 s. |
| 1.000 | 20.000 | 3.984.981 s. |
| 50 | 20.000 | 2.092.010 s. |

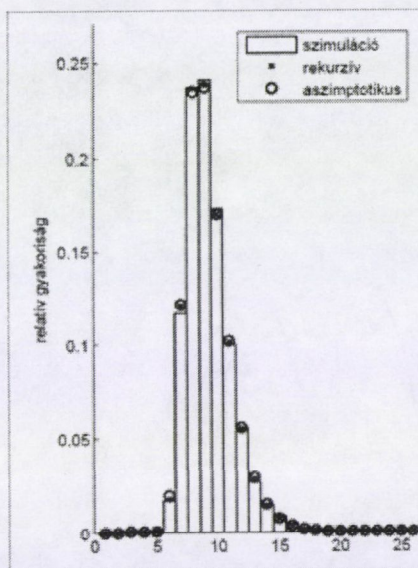
A párbaállított grafikonokon jól látszik a 2.1.2-ben leírt eredmény, miszerint R'_n eloszlása (közelítőleg) az R_n eloszlásából 1 egységgel jobbra való eltolással adódik.



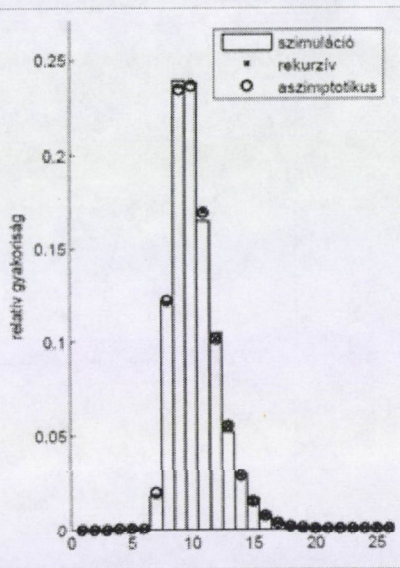
Leghosszabb fejszéria
 $p = 0,5; n = 50.$



Leghosszabb széria
 $p = 0,5; n = 50.$

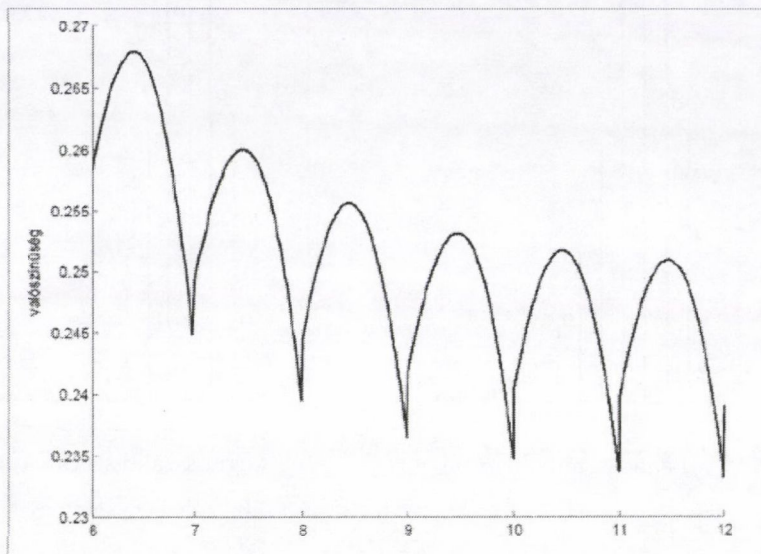


Leghosszabb fejszéria
 $p = 0,5; n = 1000.$

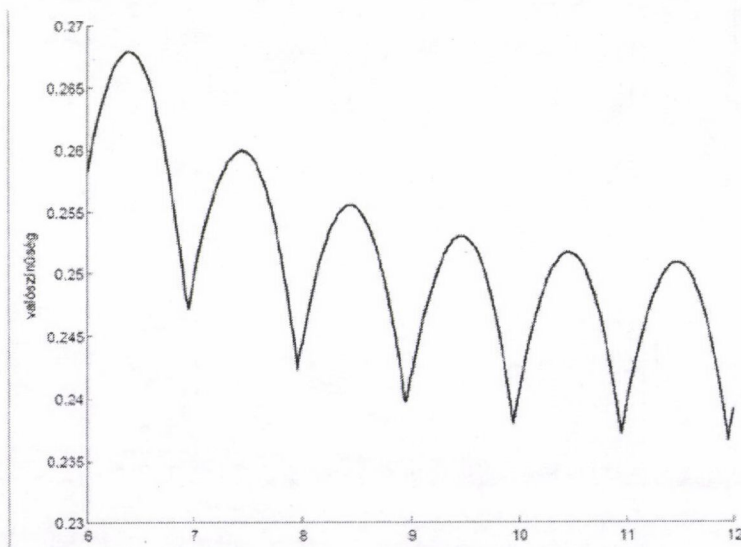


Leghosszabb széria
 $p = 0,5; n = 1000.$

A következő ábráról R_n periodikus jellegű viselkedése olvasható le, ahol a vízszintes tengelyen logaritmikus skálát alkalmaztunk, a függőleges tengelyen pedig a $P(R_n = \lfloor \log_2 n \rfloor - 1)$ értékek szerepelnek. Látható, hogy a P értékek ábrázolásakor 2 minden egész kitevős hatványa helyeken ugrás van (a rekurziós képlettel számolva).



$$P(R_n = \lfloor \log_2 n \rfloor - 1), \quad n = 2^6, \dots, 2^{12}, \quad p = 0,5$$



$$\max_k P(R_n = k), \quad n = 2^6, \dots, 2^{12}, \quad p = 0,5$$

A fenti két ábra alapján $\max_k P(R_n = k)$ értékei eltérnek $P(R_n = \lfloor \log_2 n \rfloor - 1)$ -től, azaz R_n módusza nem minden n -re lesz $\lfloor \log_2 n \rfloor - 1$.

2.2. Szabálytalan pénzérme esete

Ebben az esetben a fejdobás valószínűsége, azaz p értéke bármilyen valós szám lehet a $(0, 1)$ intervallumból. Kérdés, hogy ez a tény milyen hatással van a leghosszabb fej-, illetve leghosszabb bármilyen széria alakulására. Most nyilván nem számolhatunk a klasszikus (kedvező/összes) képlettel.

Jelölje p a fejdobás valószínűségét és $q(=1-p)$ az írás valószínűségét.

2.2.1. Leghosszabb fejszéria vizsgálata

Schilling [15] alapján tekintsük azon n hosszúságú fej-írás sorozatokat, amelyekben k db fej van. Ezek közül jelentse $C_n^{(k)}(x)$ azon sorozatok számát, amelyekben legfeljebb x fej következik egymás után. (Azaz a leghosszabb fejszéria legfeljebb x hosszúságú.) Az adott jelölésekkel az eloszlásfüggvényünk a következő lesz:

$$F_n(x) = P(R_n \leq x) = \sum_{k=0}^n C_n^{(k)}(x) p^k q^{n-k}. \quad (5)$$

2.3. ÁLLÍTÁS. (Schilling [15], 200.o.)

$$C_n^{(k)}(x) = \begin{cases} \sum_{j=0}^x C_{n-1-j}^{(k-j)}(x), & \text{ha } x < k < n, \\ \binom{n}{k}, & \text{ha } 0 \leq k \leq x, \\ 0, & \text{ha } x < k = n. \end{cases} \quad (6)$$

Bizonyítás. Ha $x < k = n$, akkor nyilvánvalóan $C_n^{(k)}(x) = 0$, hiszen ekkor az összes $(x$ -nél több) elem fej, így nincs olyan sorozat, ahol legfeljebb x fej van egymás után. Ha $0 \leq k \leq x$, akkor $C_n^{(k)}(x)$ éppen a binomiális együtthatókat adja, hiszen ez az az eset, amikor az n elem között legfeljebb x fej van, és azon eseteket kell összeszámlálni, amikor a leghosszabb fejszéria legfeljebb x . Márpedig ekkor az összes sorrend ilyen tulajdonságú. Ha $x < k < n$, akkor a (6) képlet helyességének belátásához (melyet Schilling [15]-ban nem közöl) elegendő átgondolnunk a következőket. A sorozatunk kezdődhet $j = 0, 1, 2, \dots, x$ fejjel, utána biztosan van 1 írás, majd olyan sorozat következik, ahol a maradék $n - j - 1$ elem között $k - j$ db fej van úgy, hogy a leghosszabb fejszéria legfeljebb x hosszúságú.

$$\underbrace{F \dots F}_{j \text{ db fej}} \quad I \quad \underbrace{\dots F \dots I \dots}_{n-j-1 \text{ elem, melyben } k-j \text{ db fej van úgy,}} \\ \text{hogy legfeljebb } x \text{ hosszú a fejszéria}$$

Ezek száma pedig éppen $C_{n-1-j}^{(k-j)}(x)$. Ezzel (6) helyességét beláttuk. \square

Kiszámolva $C_n^{(k)}(3)$ értékeit $n \leq 8$ esetre az alábbi értékek adódnak:

| | | | | | | | | | |
|-----|-------------------|---|---|---|---|---|---|---|---|
| 8 | 0 | | | | | | | | |
| 7 | 0 0 | | | | | | | | |
| 6 | 0 1 10 | | | | | | | | |
| 5 | 0 2 12 40 | | | | | | | | |
| 4 | 0 3 12 31 65 | | | | | | | | |
| 3 | 1 4 10 20 35 56 | | | | | | | | |
| 2 | 1 3 6 10 15 21 28 | | | | | | | | |
| 1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| k/n | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

Észrevehetjük, hogy az alsó négy sor ($k = 0, 1, 2, 3$ esetben $\binom{n}{k}$ értékek) a Pascal-háromszög részletét adja, $3 < k = n$ esetben pedig beírva a 0-kat, a táblázat többi adata a rekurziós képlet alapján számítható. Látható, hogy a fenti (6) képletet jól mutatja az ún. „hokiütő” forma. Hiszen pl. $C_7^{(5)}(3) = 2 + 3 + 4 + 3 = 12$. (Táblázatban dőlt, vastag számmal jelölve.)

Nézzük most, hogy mi lesz annak a valószínűsége, hogy n dobásból a leghosszabb fejszéria pontosan k hosszúságú? Jelöljük ezt $p(n, k)$ -val, melyre Kopocinski [8]-ban két formulát is ad. Az alábbi (7) és (8) képlet alkalmazhatóságához megjegyezzük, hogy az (5)-beli F függvényre:

$$F_n(k) = \begin{cases} 0, & \text{ha } k < 0, \\ 1, & \text{ha } k \geq n, \\ \sum_{i=0}^k p(n, i) & \text{egyébként.} \end{cases}$$

2.3. TÉTEL. (Kopocinski [8], Theorem 1, (2), (3).) Legyen $p(n, k) = 0$, ha $k < 0$, vagy $k > n$, $p(k, k) = p^k$, ha $k = 0, 1, 2, \dots$ Ekkor

$$p(n, k) = \sum_{j=0}^{k-1} p^j q p(n-j-1, k) + p^k q F_{n-k-1}(k), \quad (7)$$

$$p(n, k) = p^k q F_{n-k-1}(k) + \sum_{j=0}^{n-k-2} F_j(k-1) q^2 p^k F_{n-j-k-2}(k) + F_{n-k-1}(k-1) q p^k, \quad (8)$$

ha $n = k+1, k+2, \dots$, $k = 0, 1, 2, \dots$, és $F_n(k)$ jelöli annak a valószínűségét – az eloszlásfüggvény definíciójának megfelelően –, hogy n esetből legfeljebb k hosszúságú fejszéria adódik.

Bizonyítás. A (7) képlet helyességét vizsgáljuk teljes eseményrendszer szerinti részekre bontással aszerint, hogy az első helyeken lehet $0, 1, 2, \dots, k$ db fej. Lehet olyan, hogy az első j helyen fejet kapunk, ($0 \leq j < k$), majd jön 1 írás és az utána lévőkben van pontosan k hosszú fejszéria:

$$\underbrace{F \dots F}_{j \text{ db fej}} \quad I \quad \underbrace{\dots F \dots I \dots}_{\text{az } n-j-1 \text{ elem között}}$$

$0 \leq j \leq k-1$ pontosan k hosszú fejszéria van

Ebből az esetből adódik a (7) összeg első k tagja, azaz $\sum_{j=0}^{k-1} p^j q p(n-j-1, k)$. Vagy pedig lehet az az eset, hogy rögtön az elején adódik k hosszúságú fejszéria, utána 1 írás, majd legfeljebb k hosszúságú fejszéria:

$$\underbrace{F \dots F}_{k \text{ db fej}} \quad I \quad \underbrace{\dots F \dots I \dots}_{n-k-1 \text{ elem melyben}}$$

legfeljebb k hosszú fejszéria van

Ez pedig éppen a (7) összefüggés utolsó tagja, amivel a képlet helyességét beláttuk.

A (8) bizonyításához bontsuk fel az eseményterünket aszerint, hogy az első k hosszúságú fejszéria hol kezdődik. Lehet olyan eset, hogy rögtön az első k dobás fej, utána 1 írás, majd legfeljebb k hosszú fejszéria következik:

$$\underbrace{F \dots F}_{k \text{ db fej}} \quad I \quad \underbrace{\dots F \dots I \dots}_{n-k-1 \text{ elem között legfeljebb } k \text{ hosszú fejszéria van}}$$

Ez éppen az összeg első tagja: $p^k q F_{n-k-1}(k)$. Lehet még olyan, hogy legfeljebb $k-1$ fej van az elején, utána 1 írás kell, hogy legyen, majd következik a k db fej, utána megint 1 írás, és végül legfeljebb k db fej:

$$\underbrace{\dots F \dots I \dots}_{\text{legfeljebb } k-1 \text{ db fej}} \quad I \quad \underbrace{F \dots F}_{k \text{ db fej}} \quad I \quad \underbrace{\dots F \dots I \dots}_{\text{az } n-j-k-2 \text{ elem között}}$$

legfeljebb k db fej van

Ez éppen az összeg második része: $\sum_{j=0}^{n-k-2} F_j(k-1) q^2 p^k F_{n-j-k-2}(k)$. Vagy lehet az az eset, amikor az utolsó k db lesz fej, előtte 1 írás és a kezdőszériában legfeljebb $k-1$ fej van:

$$\underbrace{\dots F \dots I \dots}_{\text{legfeljebb } k-1 \text{ fej}} \quad I \quad \underbrace{F \dots F}_{k \text{ db fej}}$$

Ez éppen az összeg utolsó tagját adja: $F_{n-k-1}(k-1) q p^k$.

Így a (8) formula helyességét is beláttuk. (A (7) és (8) bizonyítását Kopocinski nem végzi el, csak [8] 5. oldalán útmutatást ad.) \square

Az aszimptotikus viselkedés leírását Gordon-Schilling-Waterman [7] adja a következő tétellel.

2.4. TÉTEL. (Gordon-Schilling-Waterman [7].) Legyen $\mu(n) = -\frac{\log n}{\log p}$, $q = 1-p$ és legyen W olyan, hogy teljesüljön rá: $(P(W \leq t) = \exp(-\exp(-t)))$, ekkor t -ben egyenletesen:

$$P(R_n - \mu(qn) \leq t) - P\left(\left[\frac{W}{-\log p} + \{\mu(qn)\}\right] - \{\mu(qn)\} \leq t\right) \rightarrow 0,$$

ahol $n \rightarrow \infty$. (Az előzőekhez hasonlóan $[\]$ az egészrész, $\{ \}$ pedig a törtrész jele.)

2.2.2. Leghosszabb bármilyen széria vizsgálata

Egy szabálytalan pénzérmét feldobva n -szer, jelölje R'_n a leghosszabb bármilyen széria (akár fej, akár írás) nagyságát. Az (5) képlethez hasonlóan adódik az alábbi (lásd Schilling [15], 200. oldal).

2.4. ÁLLÍTÁS. (Schilling [15].)

$$F'_n(x) = P(R'_n \leq x) = \sum_{k=0}^n \overline{C}_n^{(k)}(x) p^k q^{n-k},$$

ahol $\overline{C}_n^{(k)}(x)$ jelenti azon n hosszúságú sorozatok számát, amelyben k db fej van, a leghosszabb bármilyen széria hossza legfeljebb x , és ahol

$$\overline{C}_{m+k}^{(k)}(x) = C_{x+1}(m, k),$$

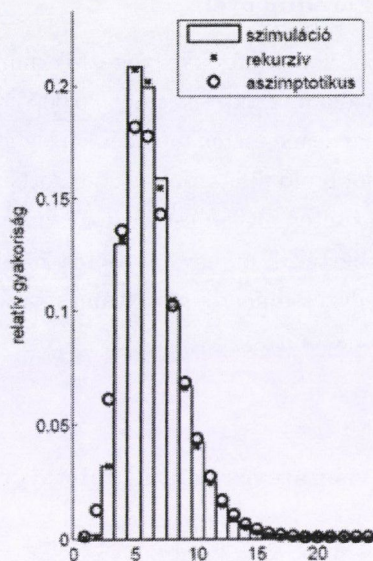
valamint a $C_x(m, k)$ mennyiségek kielégítik a (9) és (10) rekurziókat.

Itt $C_x(m, k)$ jelöli, hogy m piros és k fekete golyót visszatevés nélkül kihúзва nem lesz x hosszúságú széria. $C_t(m, k)$ értékeire Bloom [3] adott két rekurzív képletet, melyek bizonyításait a 3. fejezetben adjuk meg ((9) és (10)).

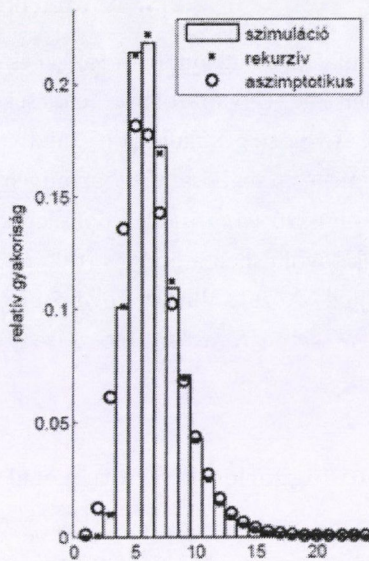
Megjegyzés. R'_n aszimptotikus viselkedését vizsgálva Muselli [9] tételét használhatjuk, melyben $V_n(p)$ jelöli annak a valószínűségét, hogy a leghosszabb széria n dobás esetén a fejekből adódik:

$$\lim_{n \rightarrow \infty} V_n(p) = \begin{cases} 0, & \text{ha } 0 \leq p < 1/2, \\ 1, & \text{ha } 1/2 < p \leq 1. \end{cases}$$

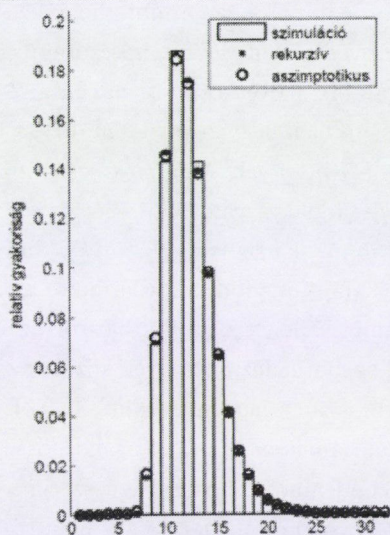
Azaz 1-hez tart annak a valószínűsége, hogy a leghosszabb széria az esélyesebb kimenetelből, azaz jelen esetben fejekből áll.



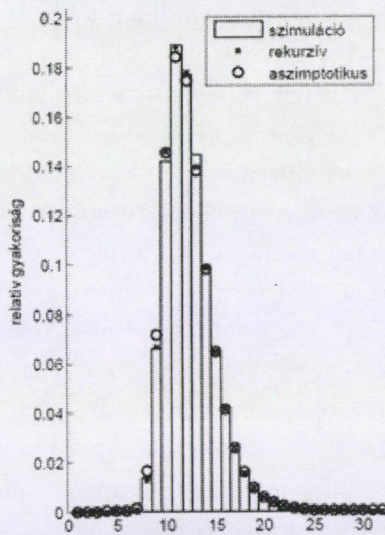
Leghosszabb fejszéria
 $p = 0,6; n = 50.$



Leghosszabb széria
 $p = 0,6; n = 50.$



Leghosszabb fejszéria
 $p = 0,6; n = 1000.$



Leghosszabb széria
 $p = 0,6; n = 1000.$

2.2.3. Numerikus eredmények

A szimulációkat ugyanolyan tárgyi és szoftver eszközökkel végeztük, mint azt a 2.1.3-ban leírtuk. Ugyanazokat a jelöléseket alkalmaztuk, és ismét 20.000 ismétlést végeztünk. Az előző grafikonokon szabálytalan érme esetén mutatjuk a leghosszabb fej (bal oldali) és a leghosszabb bármilyen (jobb oldali) szériák esetét, sorban rövid ($n = 50$) és hosszú ($n = 1000$) dobássorozat esetén. Kis n esetén az aszimptotikus eredmények távol esnek a pontos (rekurziós) értékektől, nagy n esetén közel kerülnek hozzájuk. A 2.4. állítást követő megjegyzés numerikus alátámasztását adja, hogy nagy n esetén R_n és R'_n eloszlása közel azonos.

3. Nem független kísérletsorozat (visszatevés nélküli mintavétel)

Gardner [6] könyvében szerepel az alábbi állítás: "Az 52 lapos összekevert kártyacsomagban majdnem mindig lesz 6 egyforma színű egymás után." Ahogyan [3]-ben is szerepel, elvégezve többször is ezt a kísérletet nem tapasztaltuk ezt az eredményt. 6-nál általában kevesebb elemű szériák adódtak. Csak nem volt szerencsénk, vagy a szerző tévedett?

Felvetődik az alábbi kérdés. Ha egy halmaz kétféle tulajdonságú elemet tartalmaz, az egyikből m , a másikból k db-ot, mi a valószínűsége annak, hogy az $m + k$ elemet sorban egymás után kihúzva visszatevés nélkül, lesz t hosszúságú széria, vagyis akármelyik tulajdonságú elemből legalább t következik egymás után?

Az egyszerűség kedvéért a két tulajdonságú elem legyen piros (m db) és fekete (k db), és jelöljük a keresett valószínűséget $P_t(m, k)$ -val. Meghatározásához vizsgáljuk az esemény komplementerének, vagyis annak a valószínűségét, hogy nincs t hosszúságú széria az $m + k$ elem sorozatában. Ez legyen: $\overline{P}_t(m, k)$, aminek a segítségével a $P_t(m, k) = 1 - \overline{P}_t(m, k)$ képlet alapján adódik kérdésünkre a válasz. $\overline{P}_t(m, k)$ -nek klasszikus képlettel való kiszámításához vizsgáljuk először az összes elemi esemény számát: az $m + k$ elemet kell sorbarendezni, melyek között az m db és a k db azonos típusúak. Az ilyen sorozatok száma nem más mint: $\binom{m+k}{m}$.

Ezek után a keresett hányados számlálójának meghatározásához össze kell számlálnunk azon sorozatok számát, amelyben nincs t hosszúságú széria. Jelöljük ezt $C_t(m, k)$ -val. Először számoljuk ki $C_t(m, k)$ értékeit $t = 3$ és nem túl nagy (10-nél kisebb) m és k esetén:

| $m \backslash k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------------------|---|---|---|-----------|-----------|------------|-----|-----|------|------|
| 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 1 | 2 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | 1 | 3 | 6 | 7 | 6 | 3 | 1 | 0 | 0 | 0 |
| 3 | 0 | 2 | 7 | 14 | 18 | 16 | 10 | 4 | 1 | 0 |
| 4 | 0 | 1 | 6 | 18 | 34 | 45 | 43 | 30 | 15 | 5 |
| 5 | 0 | 0 | 3 | 16 | 45 | 84 | 113 | 114 | 87 | 50 |
| 6 | 0 | 0 | 1 | 10 | 43 | 113 | 208 | 285 | 300 | 246 |
| 7 | 0 | 0 | 0 | 4 | 30 | 114 | 285 | 518 | 720 | 786 |
| 8 | 0 | 0 | 0 | 1 | 15 | 87 | 300 | 720 | 1296 | 1823 |
| 9 | 0 | 0 | 0 | 0 | 5 | 50 | 246 | 786 | 1823 | 3254 |

3.1. ÁLLÍTÁS. (Bloom [3], 369. o.) Ha $m = k = 0$, akkor definíció szerint legyen $C_t(0, 0) = 1$. Ha m vagy k negatív, akkor pedig definíció szerint $C_t(m, k) = 0$.

$$C_t(m, k) = \sum_{i=0}^{t-1} C_t(m-1, k-i) - \sum_{i=1}^{t-1} C_t(m-t, k-i) + e_t(m, k), \quad (9)$$

ahol tehát $C_t(m, k)$ jelenti az m db piros és k db fekete elem olyan sorbarendeze-seinek a számát, ahol nincs t hosszúságú széria ($t \geq 2$), valamint

$$e_t(m, k) = \begin{cases} 1, & \text{ha } m = 0 \text{ és } 0 \leq k < t, \\ -1, & \text{ha } m = t \text{ és } 0 \leq k < t, \\ 0, & \text{különben.} \end{cases}$$

Igazoljuk (9)-et, melyet Bloom [3]-ben nem végzett el.

Bizonyítás.

$m = 0$ eset.

Ha $0 \leq k < t$, akkor $C_t(0, k) = 1$, hiszen ez azt jelenti, hogy csak egyféle elem van, de kevesebb, mint a széria-hossz. Ezeket akárhogyan is húzom, nem lehet t hosszúságú széria, és mivel az azonos elemek egymás között nem megkülönböztet-hetők, ezért ez egyféle sorrendet jelent. Így a (9) képletünk a következő alakú: $1 = 0 - 0 + 1$. Például:

$$\begin{aligned} C_3(0, 2) &= C_3(-1, 2) + C_3(-1, 1) + C_3(-1, 0) - [C_3(-3, 1) + C_3(-3, 0)] + 1 = \\ &= 0 + 0 + 0 - [0 + 0] + 1 = 1. \end{aligned}$$

Ha $k \geq t$, akkor $C_t(0, k) = 0$, hiszen ekkor is csak egyféle elem van, de legalább annyi, mint a széria-hossz. Ekkor nyilván egyetlen olyan sorozat sincs, melyben ne lenne t hosszúságú széria. A (9) a következő: $0 = 0 - 0 + 0$. Például:

$$\begin{aligned} C_3(0, 4) &= C_3(-1, 4) + C_3(-1, 3) + C_3(-1, 2) - [C_3(-3, 3) + C_3(-3, 2)] + 0 = \\ &= 0 + 0 + 0 - [0 + 0] + 0 = 0. \end{aligned}$$

$0 < m < t$ esetén (9) alakja:

$$C_t(m, k) = \sum_{i=0}^{t-1} C_t(m-1, k-i) - 0 + 0.$$

Hiszen ez a következő eset:

$$\underbrace{\text{F} \dots \text{F}}_{i \text{ db fekete}} \text{ P } \underbrace{\dots \text{F} \dots \text{P} \dots}_{\substack{k-i \text{ db fekete, } m-1 \text{ db piros,} \\ \text{és közöttük nincs } t \text{ széria}}}$$

Ezek száma $C_t(m-1, k-i)$. Mivel $m < t$, így pirosból t széria nem lehet, azaz a fenti összegből nem kell levonni semmit. Példaként tekintsük:

$$\begin{aligned} C_3(2, 2) &= C_3(1, 2) + C_3(1, 1) + C_3(1, 0) - [C_3(-1, 1) + C_3(-1, 0)] + 0 = \\ &= 3 + 2 + 1 - [0 + 0] + 0 = 6. \end{aligned}$$

Ha $m = t$ és $0 \leq k < t$, ez azt jelenti, hogy az egyik fajta elemből (feketéből) kevesebb, a másiból (pirosból) pedig pontosan annyi áll rendelkezésre, mint a szériahossz. Ekkor (9) alakja a következő:

$$C_t(m, k) = \sum_{i=0}^{t-1} C_t(m-1, k-i) - \sum_{i=1}^{t-1} C_t(0, k-i) - 1.$$

Az első szumma nem t , hanem csupán $k+1$ részesetet jelent, melyek i -edik tagja olyan, hogy i db feketével kezdődik, utána 1 piros, majd $m-1$ piros és $k-i$ fekete úgy, hogy nincs t széria.

$$\underbrace{\text{F} \dots \text{F}}_{i \text{ db fekete}} \text{ P } \underbrace{\dots \text{F} \dots \text{P} \dots}_{\substack{k-i \text{ db fekete, } m-1 \text{ db piros,} \\ \text{és közöttük nincs } t \text{ széria}}}$$

Ebben egy „rossz” elem van, amikor az $m = t$ piros egymás után helyezkedik el. Mivel a második szumma értéke k , így összesen $k+1$ levonása történik. Például:

$$\begin{aligned} C_3(3, 2) &= C_3(2, 2) + C_3(2, 1) + C_3(2, 0) - [C_3(0, 1) + C_3(0, 0)] - 1 = \\ &= 6 + 3 + 1 - [1 + 1] - 1 = 7. \end{aligned}$$

Ha $m = t$ és $k \geq t$, akkor (9) a következő:

$$C_t(m, k) = \sum_{i=0}^{t-1} C_t(m-1, k-i) - \sum_{i=1}^{t-1} C_t(0, k-i) + 0.$$

Ebben $i = 0$ esetén:

$$P \quad \underbrace{\dots P \dots F \dots}_{k \text{ db fekete, } m-1 \text{ piros}} \\ \text{és nincs közöttük } t \text{ széria}$$

Ezek száma $C_t(m-1, k)$. Ebben lehetne egy „rossz” is, amikor az $m-1$ piros van elől, és a legelső pirossal t szériát alkot. De ekkor a k fekete a végén állva t szériát alkotna, vagyis ez a rossz szituáció már nincs benne $C_t(m-1, k)$ -ban.

Illetve $i = 1, 2, \dots, t-1$ esetén:

$$\underbrace{F \dots F}_{i \text{ db fekete}} \quad P \quad \underbrace{\dots F \dots P \dots}_{k-i \text{ db fekete, } m-1 \text{ db piros,}} \\ \text{és közöttük nincs } t \text{ széria}$$

Ezek száma $C_t(m-1, k-i)$. De ebben lehet egy „rossz” eset, amikor mind a $t = m$ piros egymás mellett van, így le kell vonni $C_t(0, k-i)$ mennyiséget, ami lehet 0 is. Nézzük például:

$$C_3(3, 4) = C_3(2, 4) + C_3(2, 3) + C_3(2, 2) - [C_3(0, 3) + C_3(0, 2)] + 0 = \\ = 6 + 7 + 6 - [0 + 1] + 0 = 18.$$

Ha $m > 0$ és $m > t$, akkor a következőképpen kaphatunk olyan sorozatokat, melyekben nincs t hosszúságú széria. Kezdődhet i db (t -nél kevesebb, vagyis $1 \leq i \leq t-1$) azonos elemmel (feketével), utána egy másmilyen (piros), majd ezután sincs t széria:

$$\underbrace{F \dots F}_{i \text{ db fekete}} \quad P \quad \underbrace{\dots F \dots P \dots}_{k-i \text{ db fekete, } m-1 \text{ db piros,}} \\ \text{és közöttük nincs } t \text{ széria}$$

Ezen sorozatok száma: $\sum_{i=1}^{t-1} C_t(m-1, k-i)$.

De ebben lehetnek olyanok is, ahol az 1 db piros után is pirosak vannak úgy, hogy t db piros van egymás után, és utána nincs t széria:

$$\underbrace{F \dots F}_{i \text{ db fekete}} \quad \underbrace{P \dots P}_{t \text{ db piros}} \quad \underbrace{\dots F \dots P \dots}_{k-i \text{ db fekete, } m-t \text{ db piros és}} \\ \text{nincs közöttük } t \text{ széria}$$

Ezek száma: $\sum_{i=1}^{t-1} C_t(m-t, k-i)$, amit az előző összegből le kell vonni. De ezekben lehet olyan is, hogy a t hosszú piros után is piros következik, vagyis „eltoltan” is lehet t széria. Ezek száma $\sum_{i=1}^{t-1} C_t^*(m-t, k-i)$, ahol $C_t^*(m-t, k-i)$ a pirossal kezdődő, $m-t$ pirosat és $k-i$ feketét tartalmazó olyan sorozatok száma, melyekben nincs t széria.

Nem számítottuk még be azokat az eseteket, amikor $i = 0$, vagyis pirossal kezdődik a sorozat. Ekkor az első elem piros és utána nincs t széria:

$$\begin{array}{c} P \quad \underbrace{\dots P \dots F \dots}_{k \text{ db fekete, } m-1 \text{ piros}} \\ \text{és nincs közöttük } t \text{ széria} \end{array}$$

Ezek száma: $C_t(m-1, k)$. De ebben lehetnek olyanok is, melyek t szériával kezdődnek és utána nincs t széria:

$$\begin{array}{c} \underbrace{P \dots P}_{t \text{ db piros,}} \underbrace{F \dots F}_{i \text{ db fekete}} \qquad \underbrace{P \dots F \dots P \dots}_{m-t \text{ piros, } k-i \text{ db fekete, pirossal kezdődik és}} \\ \text{nincs közöttük } t \text{ széria } (1 \leq i \leq t-1) \end{array}$$

Ezek száma: $\sum_{i=1}^{t-1} C_t^*(m-t, k-i)$, amit előző mennyiségből ($C_t(m-1, k)$ -ból) le kell vonni.

Összesítve tehát kapjuk a következőt:

$$\begin{aligned} C_t(m, k) &= \sum_{i=1}^{t-1} C_t(m-1, k-i) - \left\{ \sum_{i=1}^{t-1} C_t(m-t, k-i) - \sum_{i=1}^{t-1} C_t^*(m-t, k-i) \right\} + \\ &\quad + \left\{ C_t(m-1, k) - \sum_{i=1}^{t-1} C_t^*(m-t, k-i) \right\} + e_t(m, k). \end{aligned}$$

Ami az eredeti (9) képletünket adja. Példaként vegyük:

$$\begin{aligned} C_3(5, 2) &= C_3(4, 2) + C_3(4, 1) + C_3(4, 0) - [C_3(2, 1) + C_3(2, 0)] + 0 = \\ &= 6 + 1 + 0 - [3 + 1] + 0 = 3. \end{aligned}$$

Így tehát azon $m+k$ elemű sorozatok számát, melyben m db piros és k db fekete elemet rakunk visszatevés nélkül sorba úgy, hogy nincs t széria, valóban a (9)-es képlet szolgáltatja. \square

Az alapkérdésünkre tehát a választ a $P_t(m, k) = 1 - \frac{C_t(m, k)}{\binom{m+k}{k}}$ képletbe való behelyettesítéssel kapjuk.

A $C_t(m, k)$ értékére Bloom [3] 371. oldalán egy olyan formulát is ad, mely az m, k és t értékétől függetlenül mindig 6 tagból áll.

3.2. ÁLLÍTÁS. (Bloom [3], 371. o.) $t \geq 2$ esetén

$$\begin{aligned} C_t(m, k) &= C_t(m-1, k) + C_t(m, k-1) - C_t(m-t, k-1) - \\ &\quad - C_t(m-1, k-t) + C_t(m-t, k-t) + e_t^*(m, k), \end{aligned} \quad (10)$$

ahol

$$e_t^*(m, k) = \begin{cases} 1, & \text{ha } m = k = 0, \text{ vagy } m = k = t, \\ -1, & \text{ha } m = 0 \text{ és } k = t \text{ vagy } m = t \text{ és } k = 0, \\ 0, & \text{különben.} \end{cases}$$

Peremfeltételeink pedig ugyanazok, mint a 3.1. állításnál, vagyis, ha $m = k = 0$, akkor definíció szerint legyen $C_t(0, 0) = 1$, ha m vagy k negatív, akkor pedig definíció szerint $C_t(m, k) = 0$.

A (10) képlet helyességét Bloom [3] 371. oldalon közölttől eltérő módon az alábbiakból látjuk be.

Bizonyítás. Kezdődhet a sorozatunk 1 piros vagy 1 fekete elemmel:

$$\left. \begin{array}{c} \text{P} \underbrace{\dots \text{P} \dots \text{F} \dots}_{m-1 \text{ piros és } k \text{ fekete}} \end{array} \right\} \text{ ezek száma: } C_t(m-1, k);$$

$$\left. \begin{array}{c} \text{F} \underbrace{\dots \text{P} \dots \text{F} \dots}_{m \text{ piros és } k-1 \text{ fekete}} \end{array} \right\} \text{ ezek száma: } C_t(m, k-1).$$

Ezekből le kell vonni azokat, melyekben az első jel után is ugyanolyan következik összesen t hosszon, utána 1 másfajl és utána nincs t széria:

$$\left. \begin{array}{c} \underbrace{\text{P} \dots \text{P}}_{t \text{ piros}} \text{ F} \underbrace{\dots \text{P} \dots \text{F} \dots}_{m-t \text{ piros és } k-1 \text{ fekete}} \end{array} \right\} \text{ ezek száma: } C_t(m-t, k-1);$$

$$\left. \begin{array}{c} \underbrace{\text{F} \dots \text{F}}_{t \text{ fekete}} \text{ P} \underbrace{\dots \text{P} \dots \text{F} \dots}_{m-1 \text{ piros és } k-t \text{ fekete}} \end{array} \right\} \text{ ezek száma: } C_t(m-1, k-t).$$

De ezekben benne szerepelnek a következő sorozatok is.

A t piros, t fekete, utána pirossal kezdődően olyan $m-t$ piros és $k-t$ fekete elemű sorozat, melyben nincs t széria. Ezek száma $C_t^{(p)}(m-t, k-t)$. Illetve – fordítva – t fekete, t piros, utána feketével kezdődő olyan sorozat, melyben nincs t széria. Ezek száma $C_t^{(f)}(m-t, k-t)$. De összegük pedig éppen

$$C_t^{(p)}(m-t, k-t) + C_t^{(f)}(m-t, k-t) = C_t(m-t, k-t).$$

Összegezve tehát:

$$\begin{aligned} C_t(m, k) &= C_t(m-1, k) + C_t(m, k-1) - \\ &- \{C_t(m-t, k-1) + C_t(m-1, k-t) - C_t(m-t, k-t)\} + e_t^*(m, k). \end{aligned}$$

Vagyis valóban a (10) képletet kapjuk. \square

A szakasz elején említett kártyás példát tekintve azt találjuk, hogy míg $P_6(26, 26) = 0,464$, (annak a valószínűsége, hogy az 52 lapos kártyacsomag lapjait egymás után kirakva lesz benne egymás után 6 egyforma színű csak 0,464),

addig a $P_4(26, 26) = 0,974$. Vagyis sokkal valószínűbb a 4-es széria, így Gardner tévedett [6]-ban, amikor azt állította, hogy majdnem mindig kapunk 6-os szériát.

4. Néhány alkalmazás

Az eredmények felhasználhatók a különböző gazdasági, sorbanállási problémák, Markov-láncok (lásd pl. Samarova [14], Schuster [16] vagy Schwager [17] cikkét), tőzsdei eseménysorok vizsgálatában (lásd pl. Binswanger és Embrechts [2] cikkének 4.2. fejezetét). Az alkalmazásnak jelentős szerepe van a műszaki, pl. hidrológiai folyamatokban (lásd pl. Sen [19] cikkét) és a biológiai, pl. molekuláris biológia, DNS sorozatok vizsgálatában (lásd pl. Schilling [15] cikkét). Sok egyéb területen, például grafológiában is felhasználhatóak az eredmények (lásd pl. Arazi [1] cikkét). Révész [13] a számítástechnikában felmerülő, véletlen számsorozat számítógép általi létrehozását említi hasznos alkalmazási területként. Már több módszert is kidolgoztak, melyek véletlen fej-írás sorozatokat generálnak valamilyen módon. Nehéz azonban eldönteni, hogy a kapott véletlen sorozat valóban olyan-e, mintha igazi dobássorozatot írtunk volna. A legtöbbször attól kell tartani, hogy valamilyen – számunkra esetleg ismeretlen – periódus van a felírt sorozatban. A legtöbb statisztikai módszer nem alkalmas ennek a hibának a kiszűrésére. Egy, a leghosszabb fejszéria hosszán alapuló próba alkalmas arra, hogy a sorozatnak a periodicitás okozta nem véletlen voltát kimutassa.

Hivatkozások

- [1] ARAZI, B.: *Handwriting identification by means of run-length measurements*. IEEE Transactions on Systems, Man and Cybernetics, **7** (12), 1977, 878–881.
- [2] BINSWANGER, K. – EMBRECHTS, P.: *Longest runs in coin tossing*. Insurance Math. Econom. **15**, No. **2-3**, 1994, 139–149.
- [3] BLOOM, D. M.: *Probabilities of Clumps in a Binary Sequence*. Mathematics Magazine, **69**, No. **5**, 1996.
- [4] ERDŐS PÁL - RÉVÉSZ PÁL: *On the length of the longest head-run*. Topics in information theory (Second Colloq., Keszthely, 1975.) Colloq. Math. Soc. János Bolyai, Vol. **16**, North-Holland, Amsterdam, 1977, 219–228.
- [5] FÖLDES ANTÓNIA: *The limit distribution of the length of the longest head-run*. Period. Math. Hungar. **10**, No. **4**, 1979, 301–310.
- [6] GARDNER, M.: *aha! Gotcha*, Freeman, New York, 1982.
- [7] GORDON, L. – SCHILLING, M. F. – WATERMAN, M. S.: *An extreme value theory for long head runs*. Probab. Theory Relat. Fields, **72**, No. **2**, 1986, 279–287.

- [8] KOPOCINSKI, B.: *On the distribution of the longest succes-run in Bernoulli trials*. Roczniki Polskiego Towarzystwa Matematycznego, Seria III, Matematyka Stosowana XXXIV, 1991.
- [9] MUSELLI, M.: *Useful inequalities for the longest run distribution*. Statist. Probab. Lett. **46**, No. 3, 2000, 239–249.
- [10] PHILIPPOU, A. N. – MAKRI, F. S.: *Longest success runs and Fibonacci-type polynomials*, The Fibonacci Quarterly **23**, Nov. 1985.
- [11] RÉNYI ALFRÉD: *Probability Theory*, Akad. Kiadó, Budapest, 1970.
- [12] RÉVÉSZ PÁL: *Strong theorems on coin tossing*. Proceedings of the International Congress of Mathematicians, Helsinki, 1978.
- [13] RÉVÉSZ PÁL: *Mennyire véletlen a véletlen?* Akadémiai székfoglaló, Akadémiai Kiadó, Budapest, 1982.
- [14] SAMAROVA, S. S.: *On the asymptotic behaviour of the maximal sojourn time of an ergodic Markov chain in a fixed state*. Russian Math Surveys **35** (6), 1980, 103–104.
- [15] SCHILLING, M. F.: *The Longest Run of Heads*. The College Mathematics Journal, 1990. Vol. **21**, No. **3**
- [16] SCHUSTER, E. F.: *On overwhelming numerical evidence in the settling of Kinney's waiting time conjecture*. SIAM Journal of Statistical Computing, **6** (4), 1985, 977–982.
- [17] SCHWAGER, S. J.: *Run probabilities in sequences of Markov-dependent trials*. Journal of the American Statistical Association, **78**, 1983, 168–175.
- [18] SZÁSZNÉ SIMON JUDIT: *A sztochasztika középiskolai oktatása*. PhD értekezés, Debreceni Egyetem, 2005.
- [19] SEN, Z.: *Statistical analysis of hydrologic critical droughts*. Journal of the Hydraulics Division **106** (HY1), 1980, 99–115.

(Beérkezett: 2009. december 7.)

FAZEKAS ISTVÁN

Debreceni Egyetem

Alkalmazott Matematika és Valószínűesszámitás Tanszék

H-4010 Debrecen, Pf. 12

e-mail: fazekasi@inf.unideb.hu

KARÁCSONY ZSOLT

Miskolci Egyetem

Alkalmazott Matematikai Tanszék

H-3515 Miskolc-Egyetemváros

e-mail: matkzs@uni-miskolc.hu

LIBOR JÓZSEFNÉ

Szolnoki Főiskola

Gazdaságelemzési Módszertani Tanszék

H-5000 Szolnok, Tiszaígyeti sétány

e-mail: liborne@szolf.hu

ON LONGEST RUNS

ISTVÁN FAZEKAS, ZSOLT KARÁCSONY AND JÓZSEFNÉ LIBOR

The coin tossing experiment is considered. The length of the longest head run can be studied by asymptotic theorems (Erdős-Révész [4], Földes [5]), by recursive formulae (Schilling [15], Kopocinski [8], Bloom [3]) or by computer simulations (Binswanger-Embrechts [2]). The aim of the paper is to compare numerically the asymptotic results, the recursive formulae, and the simulation results. Moreover, we consider also the longest run (i.e. the longest pure heads or pure tails). We compare the distribution of the longest head run and that of the longest run. We consider both fair and biased coins. We also study the draw of cards without replacement. We give detailed proofs for the recursive formulae. We also touch upon a little history and applications of this topic.

EGY NEM HAGYOMÁNYOS STATISZTIKAI ELJÁRÁS BEMUTATÁSA AZ OECD PISA ADATBÁZISON - ESETTANULMÁNY

TAKÁCS SZABOLCS

A hagyományos szimulációs technikák – mint amilyenek például a bootstrap és a jackknife eljárások – csak módosítások mellett alkalmazhatóak olyan esetekben, amikor nem egyszerű véletlen mintavétel történik. A soron következőkben bemutatunk egy olyan esetet, amikor a hagyományosan alkalmazott eljárások eredményeit egy szimulációs eljárás eredményeivel vetjük össze.

Az OECD PISA felmérések során a jackknife eljárás egy módosítását alkalmazták [6] bizonyos hipotézisek kiértékelésére. A módszer elméleti hátterét [11] már vizsgálták. Az alábbiakban e módszer egy valódi nagymintás kutatás adatain történő bemutatására törekszünk – megmutatva a hagyományos és a szimulációs technikával számított eredmények közötti eltéréseket.

Az igazságot nem ismerve (lévén, nem generált adatokkal dolgozunk) próbáljuk értelmezni a különböző eljárásokból származó, eltérő eredményeket.

Megfoglalmazunk továbbá néhány észrevételt, kritikát is, amelyek az alkalmazott módszer esetén felvetődnek és az elemzéshez kapcsolódó elérhető leírásokban nem található rájuk megnyugtató válasz.

1. Bevezető

Az OECD PISA felmérés-sorozat egy 3 évente megrendezett, OECD és OECD partner országok 15 éves diákjain elvégzett oktatáspolitikai felmérés. Ennek keretében 3 tudásterületet vizsgálnak, felmérésekenként változó fókusszal. A felmérésbe kerülő iskolák száma, illetve az egy iskolából bekerülő diákok száma függ az ország nagyságától.

2000-ben a szövegértés, 2003-ban a matematikai, míg 2006-ban a természettudományos készség volt a felmérés fókuszában.

2009-ben újra a szövegértés került a figyelem középpontjába, ám a 2009-es adatok még nem nyilvánosak az elemzésünk pillanatában (vö: [18], [19], [20], [17]).

Az oktatáspolitikai felmérések során a mintaválasztás struktúrája következtében a hagyományos statisztikai eljárások helyett számítás- és így időigényes eljárásokat alkalmaznak. A mintaválasztás lényege egyfajta tömbösített rétegezésben rejlik, melynek során az adott országok nemi, települési, finanszírozási struktúráját is figyelembe veszik, és így nem egy egyszerű, véletlen mintavételezési eljárás történik.

Ráadásul a diákok súlyokat is kapnak annak elérése érdekében, hogy valóban, minden országon belül az adott ország specifikumait figyelembe véve reprezentatív mintát kapjunk.

Az elemzések során szimulációs technikákat vetnek be az országokban fellelhető, oktatás minőségét befolyásoló tényezők okozta különbségek megállapítására, azok statisztikailag szignifikáns voltának kiderítésére.

A hagyományos szimulációs technikák – mint amilyen a hagyományos bootstrap és a jackknife eljárások – csak módosítások mellett alkalmazhatóak (az alapeljárások a későbbiekben ismertetésre kerülnek). Az OECD PISA felmérések során – a reprodukálhatóság, az országoként elvégzett saját elemzések összehasonlíthatósága miatt – a jackknife eljárás egy módosítását alkalmazzák [6].

Ezt az eljárást fogjuk alkalmazni egy nem hagyományos elemzés elvégzése során [10] útmutatása alapján, a 2003-as, matematika fókuszú adatbázison.

Az adatok nyilvánosak, mind letölthetőek a [18], [19], [20] webhelyekről csak úgy, mint az elemzés elvégzéséhez szükséges eljárások és a [9] technikai segédlet.

Először a hagyományos, majd a nem hagyományos eljárások eredményeit mutatjuk be – szemléltetve azok eltérő mivoltát, ezzel némiképpen igazolva a számításigényesebb eljárás helyénvalóságát.

A két elemzésben a matematika teljesítményt fogjuk figyelemmel kísérni, és arra az egyszerű kérdésre keressük a választ, hogy a matematika teljesítmény különbözik-e a négy magyarországi képzési típusban.

Magyarországon az OECD PISA felmérésben résztvevő 15 éves diákok általános iskolába, gimnáziumba, szakközépiskolába vagy szakiskolába járnak.

A felsorolás sorrendje tudatos, az adatbázisban szereplő kódok sorrendjében történt. A kódolás sorrendje nem tükrözi a diákok a priori teljesítményét. Nem is szeretnénk, ha a teljesítmény alapján az iskolatípusok között valamifajta ordinalitás – és ezzel egyfajta megkülönböztetés – predesztinálva lenne.

Hasonló vizsgálatot végzett 2008-ban Slud és Thibaudeau [11], akik generált adatokon tesztelték a bemutatásra kerülő, nem hagyományos szimulációs eljárást.

1.1. ÁLLÍTÁS. *Slud és Thibaudeau azt tapasztalták, hogy rétegzett minták esetén, a [6] és [15] által leírt módszer eredményezi a valódi értékekhez leginkább közel álló becsléseket.*

A kísérletet generált adatokon végezték. A módszert ezen esettanulmányban egy valódi adatbázison alkalmazzuk a hozzá tartozó technikai leírások alapján.

2. A megválaszolandó kérdés, a probléma ismertetése

Az OECD PISA felmérés során hosszas elemzések és gondos előkészületek után olyan változókat hoznak létre, melyek az országoként és teljes OECD viszonylatban is jó közelítéssel normális eloszlásra vannak transzformálva.

Ez a teljesítményeket mérő változók esetén azt jelenti, hogy a teljes OECD átlag a fókusz évében olyan normális eloszlású valószínűségi változóként kezelendő, melynek várható értéke 500, szórása 100.

Megjegyzés. Az egyéb indexeket, mint amilyen pl. az ESCS vagy a SES index is, úgy transzformálják, hogy az OECD országokra vonatkoztatva standard normális eloszlást kövessenek – így a különböző országok diákjai az OECD átlagával összehasonlíthatóak lesznek ezen indexek mentén. (SES: szociokulturális háttér index).

Ez egyben azt is jelenti, hogy a vendég – partner – országok szintén az OECD szintjéhez vannak viszonyítva, de a standardizálásban ők nem vesznek részt, csak ugyanazokat a tesztek íráját és a többi országgal megegyező számítási elvek alapján határozzák meg a pontjaikat.

Egy egyszerű kérdést szeretnénk megvizsgálni, ám a választ három különböző módon fogjuk kiszámítani. A számításokhoz minden esetben az OECD PISA felmérés során elfogadhatónak ítélt és alkalmazott SPSS programcsomagot fogjuk használni.

A kérdés tehát az, hogy Magyarországon a 4 különböző képzési típusba járó 15 éves diákok átlagos teljesítménye között van-e szignifikáns különbség. Ezt az alábbi három különböző módon szeretnénk megvizsgálni.

1. Először az SPSS beépített rutinjait fogjuk alkalmazni. Értelmszerűen a fenti kérdés eldöntésére egyszempontos ANOVA-elemzést fogunk bevetni. Az első esetben az OECD PISA által használt súlyozást alkalmazzuk.
2. Második esetben módosítani fogjuk az OECD PISA felmérésben használt súlyozást és szintén az SPSS beépített eljárását alkalmazzuk a kérdés eldöntésére.
3. Harmadik esetben a felmérésben alkalmazott szimulációs eljárást fogjuk bevetni, ragaszkodva a [9]-ban megjelölt útmutatóhoz.

Ahhoz, hogy a különböző módszerek közötti különbségeket értelmezni tudjuk, illetve egyáltalán végre tudjuk hajtani az elemzéseinket, szükségünk van arra, hogy az adatbázisok felépítését megismerjük, vagy legalábbis az elemzéshez szükséges paramétereket rögzítsük.

1. Minden diák esetén a fent már említett matematika teljesítmény változó(ka)t fogjuk figyelemmel kísérni. Minden diák esetén úgynevezett plauzibilis értékeket (PV) határoztak meg, minden diákra 5 darabot. Ezt úgy kell értelmezni, mintha a diák egyetlen teljesítménye helyett 5 darab, azzal azonos eloszlású, egymástól független véletlen valószínűségi változót tekintetnénk.

Azaz:

$$PV_{M_j} \sim V(\nu_0, \sigma_0), \quad j = 1, \dots, 5,$$

és ν_0 az adott diák becsült teljesítménye, míg σ_0 az adott diák becsült teljesítményének szórása.

Ezt talán úgy lehet legkönnyebben interpretálni, hogy a tesztek során több, különböző nehézségű kérdést kapnak a diákok (mindegyik kérdés 1-1 véletlen változónak tekinthető), a teljesítmény várható értéke és szórása ezen véletlen változók együttesének figyelembevételével alakul ki [9]. A diákok teljesítményét Rasch-modell segítségével állítják elő, amely lehetővé teszi a plauzibilis értékek meghatározását [8]. E modell segítségével minden diák esetén megmondható, hogy bizonyos nehézségű feladatokat milyen valószínűséggel oldhat meg – így a diák képessége egy valószínűségi változó segítségével modellezhető. A teljesítmény meghatározásáról [9], illetve [8]-ben találhatunk információkat.

2. Minden diákot egy adott súllyal is ellátnak annak megfelelően, hogy az ország mely szempontok szerint végezte el a rétegzett mintavételt – és a diák milyen rétegnek megfelelő részpopulációt képvisel. Ennek segítségével teszik a mintát reprezentatívvá az egész országra vonatkoztatva (felsúlyozással). Itt gondolhatunk arra, hogy nemre, településtípusra, iskola típusára, méretére vonatkozó információkra, vagy a szülők iskolai végzettségére (melyet előre nem tudhatunk) kell reprezentatívvá tennünk a mintát.

Ez természetesen országonként eltérő, hiszen vannak országonként specifikus tényezők is, melyeket figyelembe kívánnak venni az adott ország felmérésében résztvevő kutatócsoportok.

Ez a felsúlyozás azt jelenti, hogy bár csak 4-5000 diák írja meg Magyarországon a tesztet, mégis úgy kezeljük a súlyozás segítségével, mintha mind a nagyjából 100000 diák megírta volna a tesztet. (Természetesen ilyenkor csak a reprezentativitás van helyreállítva, a mintánkban rejlő információ csak ilyen szempontból módosul).

Megjegyzés. A felsúlyozást a későbbiekben jól megválasztott súlyokkal próbálják ellensúlyozni. Ezért teszünk próbát arra, hogy a felsúlyozás helyett a mintán belül állítsuk helyre az arányokat anélkül, hogy az esetszámot a sokszorosára növelnénk.

3. Az előző pontban bemutatott súlyozást a második módszerhez úgy transzformáljuk, hogy az arányokat a mintán belül állítjuk helyre, így nem fogjuk a minta nagyságát mesterségesen a sokszorosára növelni – ezáltal nem csökkentjük mesterségesen a standard hibát.

Szerencsére ez a struktúra a kezdeti felmérésektől kötött, így nem okoz gondot ugyanazon programok alkalmazása a felmérések különböző időpontjaiban.

Megjegyzés. Az is világos, hogy az így felvett minta nem kezelhető egyszerű, véletlen mintaként. Így bármilyen nagy is a számunkra rendelkezésre álló adathalmaz, annak rétegzett tulajdonságait mindenképpen figyelembe kell vennünk az elemzés során.

3. A részpopulációk átlagainak összehasonlítása, hagyományos eljárás

Világos, hogy a fenti kérdés megválaszolása nem más, mint különböző csoportok esetén az átlagok meghatározása, majd azok összehasonlítása.

Ezt hagyományos ANOVA-eljárás keretei között lehet vizsgálni. Ehhez nincs másra szükség – hagyományos esetben, – mint a részminták átlagára és szórására, illetve azok elemszámára [12].

Hagyományos esetben első fázisban csak azt tudjuk megmondani, hogy a képzés típusának van-e valamilyen hatása a vizsgált változónk csoportonkénti várható értékére. Minket azonban az is érdekel, hogy ha van hatás, akkor az miben nyilvánul meg. Valójában – leginkább – erre a második kérdésre keressük a választ.

Ebben az esetben az ANOVA-elemzés egyik utóelemzésére lesz szükségünk – nevezetesen, valamely páronként is alkalmazható összehasonlításra. Ismert tény, hogy pl. páronkénti t-próbák alkalmazása nem megfelelő az elsőfajú hibák kontrolljának elvesztése miatt.

Ezért páronkénti összehasonlításra pl. a Tukey, vagy annak általánosítása-ként felfogható, robusztus Tukey–Kramer-eljárásra van szükségünk (ezen utóbbi a mintaelemszámok eltérő volta esetén is alkalmazható). Ezek minden további nélkül alkalmazhatóak a fenti esetekre attól függően – az elméleti szórások inhomogén volta esetén szabadságfok korrekciót végrehajtva. [2], [3]. Ezen módosított eljárás a Games–Howell-féle robusztus eljárás.

A vizsgált változóink normalitása részcsopontonként is biztosított, így a hagyományos eljárások kiválasztása során csak a szórások egyezésére kell tekintettel lennünk [12].

A valódi gond azonban az, hogy a mintavételi eljárásnak köszönhetően az egyedek egyáltalán nem tekinthetők függetlennek. Gondoljunk csak arra, hogy az egy iskolában (egy osztályban) tanuló diákok a hasonló (vagy azonos) tantervvvel, tanárokkal tanulnak, középiskolás korban persze eltérő érdeklődéssel – bár gyakran nagyon hasonló otthoni háttérrel.

Ráadásul azonos hatások érik őket iskolai szinten, pl. az iskola felszereltsége, iskolai közérzet, a többi diák viselkedése, tanárokkal való viszony stb.

Ahhoz, hogy ennek hatását a páronkénti összehasonlítások esetén vizsgálni tudjuk, meg kell vizsgálnunk az eljárás számítási lépéseit – itt részint a Tukey–Kramer, részint a szóráshomogenitásra robusztus Games–Howell-tesztet is vizsgálni fogjuk, mert a szórások homogenitása nem feltétlenül biztosítható. Részletezettebb leírást lásd pl. [14].

1. Jelölések: Legyenek

| | |
|-------------|---|
| n_i | az i -edik minta elemszáma, $i = 1, \dots, k$ |
| $x_{i,j}$ | az i -edik minta j -edik eleme |
| \bar{x}_i | az i -edik minta átlaga |
| s_i | az i -edik minta korrigált tapasztalati szórása |

Megjegyzés. Ezen fenti jelölések a mi példánkon az alábbiak szerint értelmezhetők:

| | |
|-------------|--|
| n_i | az i -edik iskolatípusba járó diákok létszáma, $i = 1, \dots, 4$ |
| $x_{i,j}$ | az i -edik iskolatípus j -edik diákja |
| \bar{x}_i | az i -edik iskolatípus átlaga |
| s_i | az i -edik iskolatípus korrigált tapasztalati szórása |

2. Számítsuk ki az alábbi mennyiségeket:

$$N = \sum_{i=1}^k n_i,$$

$$T_i = \sum_{j=1}^{n_i} x_{i,j} \Rightarrow \bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{i,j},$$

$$G = \sum_{i=1}^k T_i, \Rightarrow \bar{X} = \frac{G}{N},$$

$$Q_i = \sum_{j=1}^{n_i} x_{i,j}^2 - \frac{T_i^2}{n_i}.$$

3. Legyen most

$$Q_B = \sum_{i=1}^k Q_i,$$

$$Q_K = \sum_{i=1}^k n_i (\bar{x}_i - \bar{X})^2.$$

4. Legyenek továbbá

$$f_K = k - 1, \quad f_B = N - k,$$

rendre a Q_K és Q_B mennyiségekhez tartozó szabadságfokok.

5. Végző lépésként kiszámítjuk az

$$s_K^2 = \frac{Q_K}{f_K}, \quad s_B^2 = \frac{Q_B}{f_B}$$

mennyiségeket.

Megjegyzés. Az s_K^2 és s_B^2 mennyiségeket külső és belső variancia néven szokás nevezni.

3.1. KÖVETKEZMÉNY. Amennyiben a részcsoporthok várható értéke megegyezik (és az előzetesen feltett feltételek is teljesülnek), úgy az

$$F = \frac{s_K^2}{s_B^2}$$

statisztika f_K , f_B szabadságfokú F -eloszlást követ.

Ezzel még csak azt határoztuk meg, hogy van-e az iskola típusának valami hatása a diák teljesítményének szintjére. Azt azonban, hogy mely iskolatípusok teljesítenek esetleg jobban vagy rosszabbul, még nem tudjuk. Erre való pl. a Scheffé-teszt, vagy a már említett Tukey–Kramer-próba, melynek feltétele a változók normalitásán túl a részpopulációkban mért szórások egyezése [16], [22]. A szórással inhomogenitása esetén pedig a Games–Howell-teszt alkalmazandó.

A páros összehasonlítások metodikája a következő:

1. Legyenek \bar{x}_p és \bar{x}_r a két vizsgált részpopuláció átlaga, ahol $1 \leq p, r \leq k$.
2. Számítsuk ki (a fenti jelölések használatával) az alábbi statisztikát:

$$F_{p,q} = \frac{(\bar{x}_p - \bar{x}_r)^2}{(k-1) s_B^2 \left(\frac{1}{n_p} + \frac{1}{n_r} \right)}.$$

3. A normalitási és szóráshomogenitási feltétel mellett, a részpopulációk várható értékének egyezése esetén a fenti $F_{p,q}$ statisztika $f_K = k-1$ és $f_B = N-k$ szabadságfokú F eloszlást követ.

Megjegyzés. [22] szerint az eljárások egyik fontos feltétele az egyszerű, véletlen mintavételezés. Ez azt jelenti, hogy csak akkor alkalmazhatjuk ezeket az eljárásokat, ha ezen szempont szerint kellően robusztusak. Azonban, ahogy [5]-ben is láthatjuk, az eljárás csak a szórások egyezésének sérülésére robusztus, a többi feltételt nekünk kell garantálni.

A Tukey–Kramer-módszer nem sokban különbözik a fent ismertetett Scheffé-féle eljárástól.

1. Legyen továbbra is \bar{x}_p és \bar{x}_r a két vizsgált részcsoporthoz átlaga.
2. A fenti jelölések mellett számítsuk ki a

$$D_{p,r} = q_\alpha \sqrt{\frac{s_B^2}{2} \left(\frac{1}{n_p} + \frac{1}{n_r} \right)}$$

mennyiséget, ahol q_α az α szignifikancia-szinttől és k , illetve $N-k$ szabadságfoktól függő, Tukey-táblázatban megtalálható érték, lásd pl. [13].

3. Ha bármely p és r sorszámú minta átlagára teljesül, hogy

$$|\bar{x}_p - \bar{x}_r| \geq D_{p,r},$$

illetve ezzel ekvivalensen:

$$\frac{|\bar{x}_p - \bar{x}_r|}{\sqrt{\frac{s_B^2}{2} \left(\frac{1}{n_p} + \frac{1}{n_r} \right)}} \geq q_\alpha,$$

úgy azt mondhatjuk, hogy a $H_0 : \mu_p = \mu_r$ nullhipotézis elutasítható.

4. A Games–Howell-eljárást a következőképpen nyerjük: a Tukey–Kramer-eljárásban használt $D_{p,r}$ mennyiséget a szóráshomogenitási feltétel sérülése esetén módosítsuk az alábbi módon: q_α helyett q_α^f érték használandó, ahol f nem más, mint az $N - k$ szabadságfok módosítása az alábbi képlet szerint, lásd pl. [3]:

$$\begin{aligned} a &= \frac{s_p^2}{n_p}, \\ b &= \frac{s_r^2}{n_r}, \\ f &= \frac{(a+b)^2}{\frac{a^2}{n_p-1} + \frac{b^2}{n_r-1}}. \end{aligned}$$

Ez esetén kiszámítandó:

$$T_{p,r} = \frac{\bar{x}_p - \bar{x}_r}{\sqrt{\frac{a+b}{2}}}.$$

Akkor tekintjük a két várható értéket különbözőnek, ha

$$|T_{p,r}| \geq q_\alpha^f.$$

Megjegyzés. A szabadságfok kiszámításának analógiája található pl. a kétmin-tás t-próba és annak Welch-féle robusztus változatában is [12].

Megjegyzés. Az SPSS programcsomag többfajta szóráshomogenitásra nézve robusztus páronkénti összehasonlítást tartalmaz, melyek mindegyike a különbség standard hibájának egyfajta fentihez hasonló, szabadságfok korrekcióján alapul [21]. Ezek az eljárások nem mutattak érdemi különbségeket a hibabecslés konfidencia-intervallumán.

4. Nem hagyományos eljárás ismertetése

A nem hagyományos eljárás során olyan szimulációs technikát alkalmaztunk, mely egyfajta házasítása a jackknife és a bootstrap algoritmusnak. E két módszer általában jól ismert eljárások, róluk bővebben lásd pl. [6, 4, 7].

5. A jackknife eljárás

A jackknife módszer alkalmazásának feltétele mindösszesen az, hogy a mintánk független, azonos eloszlású valószínűségi változókból álljon, illetve a közös eloszlásukhoz véges szórásnégyzet tartozzon. (Ez a centrális határeloszlás tételben foglaltak teljesüléséhez szükséges).

Azt is feltételezzük, hogy a statisztikánk a megfigyeléseinkben (a valószínűségi változóinkban) szimmetrikus, tehát az argumentumok sorrendje nem befolyásolja a statisztika értékét.

Tegyük fel, hogy adott egy n elemű minta: $X_1, \dots, X_n \sim F$.

Megjegyzés. Olyan $\hat{\theta}$ statisztikákat vizsgálunk, melyek bármely n -nél kisebb elemszám esetén is értelmezettek (szükséges feltétel a kiszámíthatóság érdekében).

Bármely általános, θ paramétert becslő $\hat{\theta}$ statisztika esetén a jackknife módszerből származó becslés az alábbi alakot ölti. Jelölje $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ -et.

Tekintsük a következő n darab pszeudó-statisztikát:

$$\hat{\theta}_{(i)} = \hat{\theta}(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n).$$

Ekkor

$$\hat{\theta}_{(\bullet)} = \frac{1}{n} \sum_{i=1}^n \hat{\theta}_{(i)},$$

a keresett statisztika becslése, míg a becslés hibájának jackknife becslése:

$$\hat{\sigma}_J(\hat{\theta}) = \sqrt{\frac{n-1}{n} \sum_{i=1}^n (\hat{\theta}_{(i)} - \hat{\theta}_{(\bullet)})^2}.$$

6. A bootstrap eljárás

A bootstrap algoritmus alapját képező approximáció konvergenciájához elég-séges feltétel a mintánkat alkotó valószínűségi változók véges szórásnégyzete (közös eloszlást tételezzünk fel itt is az X_i változókra nézve). Ez a feltétel a centrális határeloszlás tétel feltételeinek teljesülése miatt szükséges is.

- (1) $\hat{\theta}(x_1, \dots, x_n)$ a statisztika értéke. (Egy adott $X_1 = x_1, \dots, X_n = x_n$ realizáció mellett).

Ekkor $\sigma(F) = \sqrt{\text{Var}_F(X_1, \dots, X_n)}$ a statisztika valódi hibája.

Ezt többnyire nehézkes zárt formában felírni.

- (2) Miután az F eloszlást nem ismerjük, ezért \hat{F} -pal, a tapasztalati eloszlásfüggvénnyel becsüljük. Ekkor persze $\hat{\sigma}_B = \sigma(\hat{F})$ becsüli $\sigma(F)$ -et. (Mint azt majd az alábbiakban látni fogjuk).

Itt csak approximációról van szó, hiszen ezt sem tudjuk zárt alakban felírni. Így tehát egy approximációs eljárást kell végrehajtanunk, mely a következő lépésekből áll.

- (i) \hat{F} meghatározása.
- (ii) F -ből független mintavétel segítségével X_1^*, \dots, X_n^* úgynevezett bootstrap minta létrehozása. Itt be kell tartanunk, hogy

$$\forall i: P(X_i^* = x_j) = \frac{1}{n}.$$

(Minden mintaelem ugyanolyan valószínűséggel veheti fel a realizációban szereplő különböző értékeket).

- (iii) $\hat{\theta}^* = \theta(X_1^*, \dots, X_n^*)$ bootstrap másolatból származó statisztika kiszámítása.
- (iv) az (ii) és (iii) lépések B számú ismétlése. Így előállítunk egy $\hat{\theta}_1^*, \dots, \hat{\theta}_B^*$ független bootstrap másolatból származó, a becslés értékére vonatkozó mintát.
- (v) $\hat{\sigma}_B$ approximáció kiszámítása az alábbi formula segítségével:

$$\hat{\sigma}_B = \sqrt{\sum_{b=1}^B \frac{(\hat{\theta}_b^* - \hat{\theta}_\bullet^*)^2}{B-1}}$$

(tapasztalati szórás), ahol

$$\hat{\theta}_\bullet^* = \frac{1}{B} \sum_{b=1}^B (\hat{\theta}_b^*).$$

Megjegyzés. Ekkor, ha $B \rightarrow \infty$, úgy a $\hat{\sigma}_B$ közelíti $\sigma(F)$ -et. B optimális megválasztásáról nincsenek különösebb viták: általában elegendő 100 és 500 közötti bootstrap minta kiszámítása.

Megjegyzés. Az eljárásokból származó, adott szintű konfidencia-intervallum könnyen meghatározható a pszeudó-statisztikákból számított, becsült paraméterek sorbarendezéséből és megfelelő méretű trimmeléséből.

Más filozófia alapján folytatható addig az (ii)–(v) lépések egymásutánja, ameddig a lépésenként kiszámított és korrigált $\hat{\theta}_\bullet^*$ valamilyen, előre meghatározott, minőséget előíró korlátnál kevesebbet változik 1 lépés alatt.

Fontos megjegyezni, hogy az approximáció konvergenciájához elégséges feltétel a véges szórásnégyzet (közös eloszlást tételezzünk fel itt is az X_i változókra nézve), mely a centrális határeloszlás tétel feltételeinek teljesülése miatt szükséges.

7. Fay eljárása és a BRR-eljárás

A jackknife eljárás szisztematikussága mellett a bootstrap eljárás véletlenszerűségét is szimuláló eljárás elméleti hátteréről [4]-ban vagy [6]-ben olvashatunk bővebben. Generált (tehát nem egy éles felmérésben végzett) adatokon végzett számításokról, azok eredményeiről [11]-ben tájékozódhatunk. Az eltérést az itt leírtaktól az adja, hogy a mi esetünkben két helyen is véletlenítés szerepel. Részint a szimulációs technika is magában hordoz egyfajta véletlenítést, részint pedig a vizsgált egyedekhez sem egyetlen mért érték tartozik, hanem 5, egymástól független, azonos paraméterekkel rendelkező, normális eloszlásból származó véletlen valószínűségi változó.

A felmérés adatbázisában alkalmazott és rögzített változók megalkotásáról részint [10] ad képet, részint pedig [23] ad tájékoztató pontokat.

Vázlatosan ismertetem a plauzibilis értékek által generált átlagok közötti különbségek standard hibáját számító algoritmust, [10] és [11] alapján.

A jól ismert jackknife módszer interpretálható úgy is, hogy az eredeti megfigyeléseket súlyozzuk: minden lépésben az egyik mintaelem súlyát 0-ra módosítjuk, a többi mintaelem súlyát változatlanul hagyjuk [15].

Fay eredeti eljárása egyik változtatásként azt mondja, hogy vegyünk egy mintaelemet $\frac{1}{2}$, míg egy bizonyos másik mintaelemet $\frac{3}{2}$ súllyal. Így annyi történik, hogy egy mintaelemet kevésbé, míg helyette egy másik mintaelemet jobban vesszünk figyelembe az egyes pszeudo-statisztikákban [6].

Az igazi különbség azonban abban rejlik, hogy bizonyos egyedek helyett csak meghatározott más egyedek léphetnek be a pszeudo-statisztikákba. Ezen csoportokat a mintavétel során feltételezett összefüggések határozzák meg (az általunk vizsgált felmérésben ezek a szempontok azok, amelyek alapján reprezentatív mintát készítettek az adott országok).

Az alaplódszert egy olyan példán ismertetem, amikor minden csoportban 2-2 egyed szerepel. Az általánosság megszorítása nélkül tehetünk ilyen egyszerűsítést. Így a jackknife módszernek egy olyan módosítását/általánosítását nyerjük, mely arányosan rétegzett mintavétel esetén is használható. (Az OECD PISA felmérésben ezeket a rétegeket és arányokat a mérést koordináló kutatók adják meg minden országra az adott sajátosságokhoz alkalmazkodva a mérést szervező OECD központnak).

Az alább ismertetésre kerülő leírás esetén a legfontosabb szempont – mely a hagyományos módszerek alkalmazását kizárja – a mintába kerülő egyedek függetlenségének hiánya.

Legyen adott H darab strátum vagy réteg. (Pl. $H = 10$ iskola). Ekkor az egy rétegbe tartozó változók helyére csak a nekik megfelelő rétegből választhatunk „helyettesítő” változót. (Azaz, csak az adott iskolák diákjai helyettesíthetők egymással).

Az egyszerűség kedvéért tegyük tehát fel, hogy minden rétegbe 2-2 egyed tartozik. Ekkor a súlyok \tilde{w} vektora: $\tilde{w} = (w_{11}, w_{12}, \dots, w_{H1}, w_{H2})$.

7.1. *Példa.* Ha a fenti 10 iskola mindegyikébe jár 20 fiú és 30 lány, akkor $\tilde{w} = (20, 30, \dots, 20, 30)$, vagy a nekik megfelelő fiú-lány létszám.

Megjegyzés. Ez még módosulhat, ha valamely oknál fogva nem veszünk valakit figyelembe (pl. a tanulási teljesítmény mérésénél a Sajátos Nevelési Igényű diákokat kihagyjuk). A végső súlyokat az egyes strátumokban ezek meghatározása után nyerjük [10].

Így jelölje a h -adik strátum i -edik esetének végső w súlyát $w_{Vhi} = f_{hi}(w)$.

8. A replikánsok súlyainak meghatározása

A BRR-eljárás ezen annyiban módosít, hogy ezt az eljárást többször, egymás után végrehajtja, mind az ismétlések miatt, mind pedig a plauzibilisek miatt [15].

Az ismétléses eljárás során a t -edik replikáns (*ismételt minta*) súlyait jelölje $w_{Vhi}^{(t)} = f_{hi}(w^{(t)})$, ahol

$$w^{(t)} = ((1 + d_{t1})w_{11}, (1 - d_{t1})w_{12}, \dots, (1 + d_{tH})w_{H1}, (1 - d_{tH})w_{H2}),$$

és itt d_{th} egy csupa 1-ből és -1 -ből álló D mátrix t -edik sorának és h -adik oszlopának eleme, ahol a mátrix sorai és oszlopai is ortogonálisak egymásra.

Megjegyzés. Az itt említésre kerülő, úgynevezett Hadamard-mátrix meghatározására [15], vagy egy program, [23] ad bővebb instrukciót. Az ortogonalitás biztosítja a minták függetlenségét, a ± 1 szorzók pedig megadják, hogy mely egyedek mely egyedek helyett kerülhetnek a replikánsokba.

Fay módosított eljárásának [6] súlyai egy korrigáló konstans szorzóban különböznek ettől, nevezetesen:

$$w^{(t)} = ((1 + d_{t1}(1 - k))\widetilde{w}_{11}, (1 - d_{t1}(1 - k))\widetilde{w}_{12}, \dots, (1 + d_{tH}(1 - k))\widetilde{w}_{H1}, (1 - d_{tH}(1 - k))\widetilde{w}_{H2}).$$

Itt $k \in [0, 1)$. Látható, hogy a BRR-eljárás a módosított Fay-eljárás speciális esete $k = 0$ esetén.

Megjegyzés. Így egyes egyedek nem kikerülnek vagy bekerülnek egy-egy replikánsba, hanem kisebb vagy nagyobb súllyal vesznek részt a paraméterek becslésében.

Amennyiben a $\theta(w_V, x)$ statisztikára vagyunk kíváncsiak, akkor a varianciák becslése:

$$(\sigma_{BRR})^2 = \frac{1}{T} \sum_{t=1}^T \left[\theta(w_V^{(t)}, x) - \theta(w_V, x) \right]^2,$$

$$(\sigma_{FAY})^2 = \frac{1}{T} \frac{1}{(1 - k)^2} \sum_{t=1}^T \left[\theta(w_V^{(t)}, x) - \theta(w_V, x) \right]^2,$$

ahol T a replikánsok (ismétlések) száma.

Megjegyzés. Hagyományos esetben a várható érték becslésének standard hibája csak a becsülni kívánt eloszlás varianciájától és az elemszámtól függ (amennyiben egyszerű véletlen mintavételezést alkalmazunk).

Az esetleges összefüggések kiküszöbölésére alkalmazható a bootstrap és a jack-knife eljárás, melyek során a standard hiba lényegében a replikánsokból számított paraméterek varianciájától és a replikánsok számától függ.

Amennyiben rétegzett mintavételről szimulációs technikát alkalmazunk, úgy a becsülni kívánt paraméterek becslésén lévő hiba most már a rétegek varianciájától és elemszámától, valamint a rétegzett mintákon külön-külön alkalmazott replikálásokban számított paraméterek varianciájától függ.

Ha minden egyes esetén még plauzibilis értékeket is számítunk, úgy a rétegzés mellett még az egyes egyedeken mért variancia is hozzáadódik a paraméter becslésének hibájához, így a teljes populációra vett becslés végső hibája három részből tevődik össze: egyik oldalról a rétegek egymáshoz viszonyított különbségeiből, másik oldalról a replikálások során számított paraméterek varianciájából, harmadrészt pedig az egyedeken lévő plauzibilisek varianciájának mértékéből.

9. A hagyományos eljárás alkalmazása

A matematika teljesítmény átlagának összehasonlítása két eltérő súlyozásra a különböző képzési típusok esetén

A hagyományos eljárás mellett számított értékek, kétfajta súlyozás mellett az alábbiak szerint alakultak:

Teljes populációra való felsúlyozás esetén:

| Képzési típus | Esetszám | Átlag | Átlag hibája | Szórás | Szórás hibája |
|----------------------|----------|---------|--------------|--------|---------------|
| 1 - Általános iskola | 6510 | 391,911 | 0,93 | 74,93 | NA |
| 2 - Gimnázium | 21027 | 551,907 | 0,37 | 72,86 | NA |
| 3 - Szakközépiskola | 41668 | 491,541 | 0,33 | 66,59 | NA |
| 4 - Szakiskola | 37837 | 405,981 | 0,43 | 62,91 | NA |

A mintán belül helyreállított arányokat alkalmazó súlyozás esetén:

| Képzési típus | Esetszám | Átlag | Átlag hibája | Szórás | Szórás hibája |
|----------------------|----------|---------|--------------|--------|---------------|
| 1 - Általános iskola | 394 | 391,911 | 3,75 | 75,06 | NA |
| 2 - Gimnázium | 1618 | 551,907 | 1,80 | 72,88 | NA |
| 3 - Szakközépiskola | 1852 | 491,541 | 1,54 | 66,60 | NA |
| 4 - Szakiskola | 901 | 405,981 | 2,09 | 62,95 | NA |

Megjegyzés. Megállapítható, hogy az átlagon lévő standard hiba - természetesen a teljes elemszám megváltozása miatt - drasztikusan megnövekedett a második súlyozás alkalmazásakor.

Megjegyzés. Az SPSS 15 nem tartalmaz olyan rutint, mely pl. a szórás konfidenca-intervallumát meghatározná.

Az átlagét az ismert $\frac{\sigma}{\sqrt{n}}$ képlet segítségével számított standard hibával meghatározhatjuk, míg a leíró statisztikák menüpontban opcionálisan a ferdeség és csúcsosság paraméterek mellett azok standard hibáját is kiszámítja.

Az is megfigyelhető, hogy a második súlyozás esetén a szórás néhány tizeddel eltérő értéket vett fel (ez teljesen elfogadható, ha figyelembe vesszük a korrigált tapasztalati szórás számítását).

Ezek után vizsgáljuk meg, hogy a hagyományos eljárás, illetve annak szórás-homogenitást nem igénylő, robusztus, páronkénti összehasonlítást elvégző számításai mit eredményeztek a kétféle súlyozás esetén. Azért a robusztus eljárást alkalmazzuk, mert a Levene-próba alapján sérül a szórások egyenlőségét igénylő szórás-homogenitási feltételünk.

Először ismételten a felsúlyozott, majd az arányaiban helyreállított súlyozás eredményeit ismertetem.

Eredmények felsúlyozás esetén:

| Képzési típus | Átlagok különbsége | Különbség hibája | Konf. iv. alja | Konf. iv. teteje |
|---------------|--------------------|------------------|----------------|------------------|
| 1 - 2 | -159,996 | 1,00 | -162,56 | -157,42 |
| 1 - 3 | -99,63 | 0,98 | -102,16 | -97,1 |
| 1 - 4 | -14,071 | 1,03 | -16,7 | -11,44 |
| 2 - 3 | 60,366 | 0,50 | 59,09 | 61,64 |
| 2 - 4 | 145,926 | 0,57 | 144,45 | 147,4 |
| 3 - 4 | 85,56 | 0,54 | 84,16 | 86,95 |

Eredmények arányosított, második számú súlyozás esetén:

| Képzési típus | Átlagok különbsége | Különbség hibája | Konf. iv. alja | Konf. iv. teteje |
|---------------|--------------------|------------------|----------------|------------------|
| 1 - 2 | -159,996 | 4,75 | -172,26 | -147,73 |
| 1 - 3 | -99,63 | 4,67 | -111,69 | -87,57 |
| 1 - 4 | -14,071 | 4,87 | -26,62 | -1,52 |
| 2 - 3 | 60,366 | 2,35 | 54,31 | 66,42 |
| 2 - 4 | 145,926 | 2,72 | 138,94 | 152,91 |
| 3 - 4 | 85,56 | 2,57 | 78,94 | 92,18 |

Jól látható, hogy mindkét esetben, 95%-os szignifikancia-szint mellett a különbségek mindenhol eltérnek 0-tól, azaz a vizsgált várható értékek egyenlősége elvetethető. Még a leginkább közel lévő, általános iskola és szakiskola esetén is szignifikáns különbség adódott.

10. A nem hagyományos eljárás alkalmazása

A matematika teljesítmény átlagának összehasonlítása a különböző képzési típusok esetén

Ebben az esetben is a 4 különböző képzési típusban vizsgáltuk meg a matematika teljesítmény átlagát és szórását. Az átlag esetén képes az SPSS hibát számítani – ezt fel is tüntetjük – azonban egy lényeges különbség rögtön látható lesz. Az előzőekkel ellentétben a teljesítmény szórása esetén az SPSS nem képes standard hibát meghatározni (vagy akár konfidencia-intervallumot). A szimulációs eljárások ebben (is) segítségünkre lehetnek, hiszen a szimulációs technikákkal az eloszlás bármely paraméterének standard hibája becsülhető, illetve bármely paraméterre konfidencia-intervallum illeszthető.

Az átlagok esetén itt is páronkénti összehasonlítást alkalmazó ANOVA-elemzést alkalmaztunk. Így megállapíthatjuk, hogy a különböző képzési típusok esetén valóban vannak-e teljesítménybeli eltérések.

A Fay-féle módosított BRR-szimulációs eljárással számított értékek az alábbiak lettek (a súlyozás készítésekor figyelembe veszik tehát azt, hogy az egy iskolában tanuló diákok nem tekinthetők egymástól függetlennek):

| Képzési típus | Esetszám | Átlag | Átlag hibája | Szórás | Szórás hibája |
|----------------------|----------|---------|--------------|--------|---------------|
| 1 - Általános iskola | 6510 | 391,911 | 7,16 | 82,67 | 4,54 |
| 2 - Gimnázium | 21027 | 551,907 | 5,00 | 77,19 | 1,91 |
| 3 - Szakközépiskola | 41668 | 491,541 | 4,61 | 71,18 | 1,95 |
| 4 - Szakiskola | 37837 | 405,981 | 6,41 | 68,50 | 2,2 |

Megjegyzés. A csoportonkénti szórások nem egyeznek meg azzal, amit a hagyományos eljárásokkal, a plauzibilis értékek helyett azok átlagával számítottunk. Ez abból fakad, hogy a diákonkénti átlagot (elvárt teljesítményt) még tudtuk garantálni, azonban a plauzibilis értékek átlagának alkalmazásával az egy diákon lévő bizonytalanságot megszüntettük – így ezt a varianciából is eltüntettük.

A csoportok közötti különbségek és azok standard hibája, illetve az ezekből számított 95%-os konfidencia-intervallumok az alábbi táblázatban láthatóak. A kódokat az előző táblázatokban már ismertettük.

| Képzési típus | Átlagok különbsége | Különbség hibája | Konf. iv. alja | Konf. iv. teteje |
|---------------|--------------------|------------------|----------------|------------------|
| 1 - 2 | -159,996 | 8,39 | -176,44 | -143,55 |
| 1 - 3 | -99,63 | 8,45 | -116,2 | -83,06 |
| 1 - 4 | -14,071 | 9,58 | -32,86 | 4,71 |
| 2 - 3 | 60,366 | 6,84 | 46,95 | 73,78 |
| 2 - 4 | 145,926 | 8,17 | 129,91 | 161,94 |
| 3 - 4 | 85,56 | 7,8 | 70,28 | 100,84 |

Megfigyelhető a különbség a hagyományos eljárás és a nem hagyományos, szimulációs eljárás eredményei között.

Míg a hagyományos esetben, szóráshomogenitást nem igénylő robusztusabb változata esetén is, mindkét súlyozásra azt az eredményt kaptuk, hogy a különböző iskolatípusok esetén a matematika teljesítmény várható értéke szignifikánsan eltér, addig a szimulációs technika azt mutatja, hogy bár vannak szignifikáns eltérések, de kivétel mégis akad.

A 15 éves diákok közül azok, akik szakiskolákba vagy általános iskolákba járnak, a szimulációs módszer alapján nem különböztek szignifikáns módon az átlagos teljesítményüket tekintve. Megállapítható, hogy a gimnazisták messze a többi iskolatípus felett teljesítettek, míg a szakközépiskolások a két teljesítmény között helyezkednek el, valamivel az OECD országok átlaga alatt.

11. Összegés: a különböző eljárások eredményeinek összehasonlítása

Az általános formulában használt két eljárás (az egyik az eredeti súlyozással, felfűjt minta esetén, a másik az eredeti súlyozás mértékét megtartva, eredeti mintanagyságon belül dolgozva) alul becsüli azt a hibát, amit szimulációs eljárás segítségével nyerünk.

Ez olyan szempontból mindenképpen eltérő eredményre vezet, hogy a kisebb standard hiba következtében a különbségek hamarabb válnak szignifikánssá.

A szimulációs eljárás [11] alapján elméletileg bizonyítottan jobb becslést nyújt számunkra a valós standard hibát illetően rétegzett mintavételezés esetén, hiszen figyelembe veszi azt, hogy a mintánk korántsem rendelkezik azzal a függetlenséggel, mely a megszokott, $\frac{\sigma}{\sqrt{n}}$ standard hiba képletének alkalmazásához szükséges.

Bár e módosított eljárások némiképpen számításigényesebbek, még ekkora minták esetén sem okoznak érdemi számításidő növekedést. Az alkalmazásukból származó esetlegesen pontosabb információ mindenképpen megéri a ráfordított időt. Azonban csak valószínűsíteni tudjuk azt, hogy a szimulált eredmények megbízhatóbbak erre a felmérésre, mint a hagyományos eljárások számításai.

Mindenképpen meg kell azonban fogalmazni legalább egy kritikát is ezzel a módszerrel kapcsolatban: az alkalmazásban található BRR-technika hivatott kiküszöbölni a mintába kerülő egyedek közötti összefüggéseket, melyek az egy strátumba való tartozás miatt fellépnek. Nem világos, hogy az adatbázisban található BRR-technika hogyan tudja kiküszöbölni azt a problémát, hogy az egy strátumba kerülő egyedek közötti összefüggések változónként eltérőek – az eljárásban azonban, változótól függetlenül mindig ugyanazokat replikáns variációkat alkalmazzuk.

Egyszerűsítési okokból nem a bootstrap eljárást alkalmazzák ebben az elemzésben, az azonban nem látszik tisztázottnak, hogy milyen módon lehet visszanyerni ebben a módszerben a bootstrap algoritmus empirikus eloszlást alkalmazó erősségét, mely strátumonként, minden változóra más és más. Kisebb elemszám

mellett a bootstrap algoritmus alkalmazhatatlan lenne, azonban az OECD PISA adatbázis kellő elemszámmal rendelkezik ahhoz, hogy ez a probléma ne merüljön fel. [1]-ben a jackknife algoritmus különböző variánsaival összehasonlítják, azonban a jackknife algoritmus semmilyen módon nem veszi figyelembe a vizsgált változók empirikus eloszlását.

Hivatkozások

- [1] BRICK, J. M., MORGANSTEIN, D. AND VALLIANT, R.: *Analysis of Complex Sample Data Using Replication*, WESTAT, (2000).
- [2] DUNNETT, CH. W.: *Pairwise Multiple Comparison in the Homogeneous Variance, Unequal Sample Size Case*, Journal of the American Statistical Association **75**, (1980) 789–795.
- [3] DUNNETT, CH. W.: *Pairwise Multiple Comparison in the Unequal Variance Case*, Journal of the American Statistical Association **75**, (1980) 796–800.
- [4] EFRON, B. AND GONG, G. (1983): *A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation*, The American Statistician **37**, 36–48.
- [5] HUZSVAI LÁSZLÓ: *Biometriai módszerek az SPSS-ben, SPSS alkalmazások*, Debreceni Egyetem, Mezőgazdaságtudományi Kar, Debrecen, (2004) 53–36.
- [6] JUDKINS, D. R.: *Fay's method for Variance Estimation*, Journal of Official Statistics Vol. **6** No. **3**, (1990) 223–239.
- [7] KÁRÁSZ JUDIT (2004): *A bootstrap algoritmus és a jackknife módszer*, Szakdolgozat, ELTE
- [8] MOLNÁR GYÖNGYVÉR (2005): *Az objektív mérés lehetősége: a Rasch-modell*, Iskolakultúra, 2005/3, 71–80.
- [9] *PISA2006 Technical Report*, OECD, (2009).
- [10] *PISA2006 Technical Report*, OECD, (2009) 188–192.
- [11] SLUD, E. V. AND THIBAUDEAU, Y.: *BRR versus Inclusion-Probability Formulas for Variances of Nonresponse Adjusted Survey Estimates*, AMSTAT, Section on Survey Research Methods, JSM (2008) 2057–2064.
- [12] VARGHA ANDRÁS: *Matematikai Statisztika pszichológiai, nyelvészeti és biológiai alkalmazásokkal (2. kiadás)*, Pólya Kiadó, Budapest, (2007).
- [13] VARGHA ANDRÁS: *Statisztikai táblázatok*, Tankönyvkiadó, (1983), XXXIII. táblázat.
- [14] VARGHA ANDRÁS: *Pszichológiai statisztika gyakorlat II.*, Tankönyvkiadó, (1981) 112–124.
- [15] WOLTER, K. M.: *Introduction to variance estimation*, Springer, New York, (1985).
- [16] <http://people.richland.edu/james/lecture/m170/ch13-dif.html>
- [17] www.pisa.oecd.org
- [18] <http://pisa2000.acer.edu.au>
- [19] <http://pisa2003.acer.edu.au>

- [20] <http://pisa2006.acer.edu.au>
- [21] <http://www.psychology.nottingham.ac.uk/staff/pal/stats/C82MST/contrasts.pdf>
- [22] <http://xenia.sote.hu/hu/biosci/docs/biometr/lecture/anova1.html>
- [23] <http://www.westat.com>

(Beérkezett: 2010. február 1.)

TAKÁCS SZABOLCS

Károli Gáspár Református Egyetem, Bölcsésztudományi Kar
Pszichológiai Intézet, Általános lélektani és módszertani tanszék
1037, Budapest, Bécsi út 324., 5. épület, fszt.
E-mail: tretarkhon@gmail.com

A NON-TRADITIONAL STATISTICAL METHOD IN THE OECD PISA DATABASE CASE-STUDY

SZABOLCS TAKÁCS

When we haven't got a simple random sample the original simulation technics, like the bootstrap and the jackknife method are usable only with changes. In the following article we show a case when we compare the results of the traditional method and the result of a simulation algorithm.

One of the jackknife's modification [6] is used for some hypothesis in the OECD PISA survey. The theorems of the applied method are introduced in [6], [11]. Now we would like to show these algorithms in a large sample survey and we show the difference between the empirical results of the traditional and the simulation method.

Because we don't use generated database we don't know the truth. Without this information we would like to understand the different results of the algorithms.

We put some questions and critics to these algorithms which are relevant and we can't find reassuring answers in the connected bibliography of the survey.

TERMÉSZETES VÍZFOLYÁSOK ÖNTÖZŐVÍZ KÉSZLETÉNEK MEGHATÁROZÁSÁT CÉLZÓ MATEMATIKAI MODELLEK

PRÉKOPA ANDRÁS, SZÁNTAI TAMÁS, ZSUFFA ISTVÁN

A dolgozat első két szakaszában a természetes vízfolyások vízhozam intenzitását vizsgáljuk valószínűségelméleti alapon. Megmutatjuk, hogy bizonyos, reális feltételek mellett ez a valószínűségeloszlás gamma típusú. A harmadik szakaszban a vízkivételi műtárgy optimális kapacitásának meghatározását célzó modellt fogalmazunk meg, midőn víztározásra nincs lehetőség, vagy azt nem szándékozunk igénybe venni. A negyedik szakaszban a modell egy optimális tározókapacitás meghatározását és az optimális vízhasznosítást célozza, miközben a vízigények nagy valószínűséggel való teljesítését előírjuk. Végül az utolsó, ötödik szakaszban az előbb említett, megbízhatósági feltételhez hozzávesszük még azt, hogy a vízigények legfeljebb adott számú napon nem teljesülnek és egy vízkivételi műtárgy kapacitásának költségét minimalizáljuk.

1. Poisson-folyamat által származtatott másodlagos sztochasztikus folyamatok

A Poisson-típusú sztochasztikus folyamat gyakran előfordul a hidrológiában. Jelen esetben az esőzések időpontjainak egymásutánjáról tételezzük fel, hogy Poisson-folyamatot alkot. Ezen azt értjük, hogy ha $\xi(I)$ jelöli egy I időintervallumban történő esőzések (véletlen) számát, akkor

- a) minden olyan I_1, \dots, I_n intervallumrendszer esetén, melyben bármely két intervallumnak nincs közös belső pontja, a $\xi(I_1), \dots, \xi(I_n)$ valószínűségi változók függetlenek,
- b) $\xi(I)$ Poisson-eloszlású $\lambda(I)$ paraméterrel, ahol $\lambda(I) \geq 0$.

Egy Poisson-folyamat által származtatott másodlagos folyamat azt jelenti, hogy a Poisson-folyamat véletlen eseményeihez, az adott esetben az esőzési időpontokhoz tartozik egy véletlen másodlagos jelenség, az adott esetben ez egy véletlen árhullám. A másodlagos jelenségek eseményterét jelölje Y . Ebben értelmezve van egy vagy több valószínűségi mérték.

A másodlagos folyamatok tárgyalására igen alkalmas az ún. szorzattér módszer [3]. Ez abban áll, hogy a másodlagos folyamatot a (t, y) elempárok halmazában,

másszóval a $T \times Y$ szorzattérben tekintjük, ahol T az időtengely egy részhalmaza, ennek eleme t . A másodlagos folyamat egy speciális lefutása, realizációja egy véletlen pontrendszerként jelent a $T \times Y$ térben. Valóban, ha $\dots, t_{-1}, t_0, t_1, t_2, \dots$ alkotják a Poisson-típusú véletlen eseményfolyamatot és $\dots, y_{-1}, y_0, y_1, y_2, \dots$ a megfelelő másodlagos jelenségek sorozata, akkor a másodlagos folyamat adott realizációja jellemezhető a $T \times Y$ térbeli

$$\dots, (t_{-1}, y_{-1}), (t_0, y_0), (t_1, y_1), (t_2, y_2), \dots$$

véletlen pontrendszerrel.

A másodlagos folyamatok szorzattérszerű elméletének alaptétele a következőt mondja ki [3].

Ha a Poisson-folyamat különböző eseményeihez tartozó másodlagos jelenségek Y térből való kiválasztása egymástól független és azonos eloszlású, μ valószínűségi mértékkel, akkor a $T \times Y$ térben elhelyezkedő véletlen pontrendszer szintén Poisson-típusú $\lambda \times \mu$ paraméter-mértékkel.

Ha a Poisson-folyamat különböző időpontokban bekövetkező eseményeihez tartozó másodlagos jelenségek függetlenek, de a t időpontban bekövetkezett eseményhez tartozó másodlagos jelenség valószínűségeloszlása függ a t -től, tehát az árhullám levonulása függ attól, hogy az árhullám kiváltása mely időpontban történt, akkor a μ mérték helyett μ_t mértékekről kell beszélnünk, és a szorzattérbeli Poisson-típusú véletlen pontrendszer D halmazhoz tartozó paraméter-mértékét az alábbi integrál származtatja

$$\int_C \mu_t(D_t) \lambda(dt), \quad (1)$$

ahol C a D halmaz vetülete a T halmazra és D_t pedig a D halmaz közös része a $T \times Y$ tér ama részhalmazával, melyen t állandó, vagyis $D_t = \{y \mid (t, y) \in D\}$.

A $T \times Y$ szorzattér D halmazába eső véletlen pontok számát $\eta(D)$ jelöli. Az (1) integrál tehát $E(\eta(D))$ -vel egyenlő.

A továbbiakban az egyszerűség kedvéért feltesszük, hogy μ_t független t -től.

A másodlagos folyamatok elméletének egy másfajta tárgyalása található a [10] dolgozatban.

2. Az öntözővízkészlet meghatározása a másodlagos folyamatok elméletére támaszkodó modell alapján

Az árhullám-intenzitás időbeli lefutását jellemezze egy $f(t, \kappa)$ függvény, ahol κ valószínűségi változó, esetleg valószínűségi vektorváltozó. Ennek egy lehetséges változata a következő

$$f(t, \kappa) = \kappa t^{\alpha-1} e^{-\beta t}, \quad t \geq 0. \quad (2)$$

A t_i időpontban történt esőzés által kiváltott árhullám az

$$f(t - t_i, \kappa_i), \quad t \geq t_i \quad (3)$$

függvénynek megfelelően vonul le, ahol a különböző i indexekhez tartozó κ_i valószínűségi változók függetlenek. A vízhozamintenzitást a (3) függvények szuperpozíciója, tehát a

$$\sum_{t_i \leq t} f(t - t_i, \kappa_i) = \eta_t$$

függvény írja le. Meghatározzuk az η_t valószínűségi változó eloszlását a (2) függvény esetében.

Alaptételünkből következik, hogy a t időpontban az (a, b) határok között futó intenzitásgörbék száma Poisson-eloszlást követ, melynek paraméterét az alábbi integrál adja meg

$$\int_{-\infty}^t P(a \leq \kappa(t-x)^{\alpha-1} e^{-\beta(t-x)} \leq b) \lambda(dx). \quad (4)$$

Erre vonatkozólag az $a = y$, $b = y + dy$ esetben az alábbi eredményt kapjuk, feltételezve még, hogy $\lambda(dx) = \lambda dx$, ahol $\lambda \geq 0$ állandó:

$$\begin{aligned} & \int_{-\infty}^t P(y \leq \kappa(t-x)^{\alpha-1} e^{-\beta(t-x)} \leq y + dy) \lambda dx \\ &= \int_0^{\infty} P(y \leq \kappa v^{\alpha-1} e^{-\beta v} \leq y + dy) \lambda dv \\ &= \int_0^{\infty} \frac{d}{dy} \left[1 - e^{-y \delta e^{\beta v} v^{1-\alpha}} \right] dy \lambda dv \\ &= \lambda \int_0^{\infty} \delta e^{\beta v} v^{1-\alpha} e^{-y \delta e^{\beta v} v^{1-\alpha}} dv dy, \end{aligned}$$

ahol κ exponenciális eloszlású $1/\delta$ várható értékkel. Az exponenciális eloszlás feltételezése nem lényeges, más eloszlást is választhatunk.

A t időpontbeli vízhozamintenzitás eloszlásfüggvénye a következő módon kapható meg. Jelölje $\eta(I)$ a t időpontban az I intervallumban elhelyezkedő egyedi intenzitásgörbék számát. Ekkor az előbbiek szerint $\eta(I)$ Poisson-eloszlású (4) paraméterrel az $I = (a, b)$ esetben. Eszerint a keresett valószínűségeloszlás karakterisztikus függvénye:

$$e^{\int_0^{\infty} (e^{iuy} - 1) E(\eta(dy))} = e^{\lambda \int_0^{\infty} (e^{iuy} - 1) \left[\int_0^{\infty} \delta e^{\beta v} v^{1-\alpha} e^{-y \delta e^{\beta v} v^{1-\alpha}} dv \right] dy}. \quad (5)$$

Az $\alpha = 1$ esetben eredményként az alábbi adódik:

$$e^{\lambda \int_0^{\infty} (e^{iuy} - 1) \frac{e^{-\delta y}}{\beta y} dy},$$

ami egy gamma-eloszlás karakterisztikus függvénye. Ugyanis, ha $\alpha = 1$, akkor (5) így folytatható:

$$e^{\lambda \int_0^{\infty} (e^{iuy} - 1) \left[\int_0^{\infty} \delta e^{\beta v} e^{-y\delta e^{\beta v}} dv \right] dy} = e^{\lambda \int_0^{\infty} (e^{iuy} - 1) \frac{1}{\beta y} e^{-\delta y} dy}. \quad (6)$$

A gamma-eloszlás karakterisztikus függvényének (6) alakban való előállítását illetően lásd a [2] könyv 92. oldalát.

Az ebben a szakaszban alkalmazott megfontolások átvihetők bonyolultabb $f(t, \kappa)$ árhullámfüggvények esetére is. Az eredmény azonban nem feltétlenül képletszerű, de mindenképpen numerikusan nyerhető. Eredményként η_t eloszlását és a vízhiányos időszak átlagos hosszúságát tudjuk nyújtani.

3. Sztochasztikus programozási modell, melyben a vízkivételre szolgáló készülék kapacitása a meghatározandó számérték

Egymás utáni időszakokat (periódusokat) tekintünk és bevezetjük a következő jelöléseket:

| | |
|----------------------------|--|
| η_k | a szolgáltatandó öntözővíz iránti igény nagysága a k -adik periódusban: $\eta_k = h_k - \gamma_k$, ahol h_k állandó és az összes vízigényt jelenti, γ_k pedig a csapadék mennyisége a k -adik periódusban, |
| ξ_k | a vízhozam a k -adik periódusban, |
| m | a készülék kapacitása, |
| M | az m kapacitást ésszerűen korlátozó felső határ, |
| $p(m)$ | a készülék ára a kapacitás függvényében, |
| $\zeta_k = \min(m, \xi_k)$ | a k -adik periódusban kiszolgáltatható vízmennyiség, |
| c_k | egységnyi víz haszna a k -adik periódusban, |
| K | a periódusok száma, |
| N | az évek száma, |
| $p > 0$ | N éven át konstansnak feltételezett inflációs ráta. |

Tételezzük fel, hogy a k -adik periódusban vízhiány esetén a hiányzó víz mennyiségével arányos kár keletkezik. Ez a feltételezés egyébként implicite már szerepelt c_k bevezetésekor, mert a c_k számnak csak az előbbi feltételezés mellett van értelme; c_k egyébként nem más, mint az említett pozitív arányossági tényező. A tárgyalandó feladat a nemlineárisan növekvő büntetés esetére is megfogalmazható.

A k -adik periódusban keletkezett véletlen nagyságú kárt az alábbi valószínűségi változó szolgáltatja

$$\chi_k = c_k [(\eta_k - \zeta_k)]_+ = \begin{cases} c_k(\eta_k - \zeta_k), & \text{ha } \eta_k > \zeta_k, \\ 0, & \text{egyébként.} \end{cases}$$

K egymásutáni periódust tekintve, az összes bekövetkező kár várható értéke $\sum_{k=1}^K E(\chi_k)$ lesz. Ha a várható összes kár nagyságát a tárgyévra és az azt követő N egymás utáni évre összegezve szeretnénk minimalizálni, akkor az okozott károk várható értékét jelenértéken tekintve, az alábbi optimalizálási feladatot kell megoldani:

$$\min \left[p(m) + \sum_{i=0}^N \left(\sum_{k=1}^K E(\chi_k) \right) \frac{1}{(1+p)^i} \right], \quad \text{feltéve, hogy } 0 \leq m \leq M \quad (7)$$

A (7) feladat egyváltozós optimalizálási feladat, a $[0, M]$ intervallumon keressük a célfüggvény minimumát. Megmutatjuk, hogy a célfüggvényben szereplő összeg az m változó konvex függvénye. Elegendő a konvexitást egy tagra megmutatni. Jelölje G_k az η_k , F_k a ζ_k valószínűségi változó eloszlásfüggvényét, f_k pedig az ehhez tartozó sűrűségfüggvényt. Ekkor a ζ_k valószínűségi változó definíciója szerint:

$$\begin{aligned} \frac{1}{c_k} E(\chi_k) &= E([\eta_k - \zeta_k]_+) \\ &= \int_0^m E([\eta_k - z]_+) f_k(z) dz + \int_m^\infty E([\eta_k - m]_+) f_k(z) dz \\ &= \int_0^m \left(\int_z^\infty (1 - G_k(x)) dx \right) f_k(z) dz + \int_m^\infty (1 - G_k(x)) dx (1 - F_k(m)) \\ &= \int_0^m \int_0^\infty (1 - G_k(y+z)) f_k(z) dy dz + \int_m^\infty (1 - G_k(x)) dx (1 - F_k(m)). \end{aligned} \quad (8)$$

A levezetésben felhasználtuk, hogy ha egy ξ valószínűségi változónak $f(x)$ a sűrűségfüggvénye, $F(x)$ az eloszlásfüggvénye és létezik a várható értéke, akkor parciális integrálással könnyen igazolható, hogy tetszőleges z valós számra

$$E([\xi - z]_+) = \int_z^\infty (x - z) f(x) dx = \int_z^\infty (1 - F(x)) dx.$$

A (8) formula m változó szerinti kétszeres deriválásával meggyőződhetünk arról, hogy $(1/c_k)E(\chi_k)$ konvex függvény. Minthogy $c_k > 0$, $k = 1, \dots, N$, következik,

hogy $E(\chi_k)$ és ezek összege is konvex. $p > 0$ miatt pedig az N évre összegzett és jelenértékre hozott várható kár mennyisége is az m változó konvex függvénye. Ha $p(m)$ is konvex, akkor az egész célfüggvény konvex. Ha $p(m)$ nem konvex, akkor a célfüggvény konvexitása nem bizonyítható, azonban bizonyos esetekben ettől még lehet akár konvex is, amint az a példánk esetében is látható. Az optimalizálás viszonylag egyszerűen elvégezhető. A vízhozamok eloszlására választhatjuk a gamma, a vízigények eloszlására pedig a normális eloszlást.

A (7) modellnek több variánsa fogalmazható meg. A további lehetőségek lényegében a vízszolgáltatás folyamatosságának előírt megbízhatósági szintjét tartalmazzák valamilyen formában.

A (7) modellt azon a numerikus példán szemléltetjük, amelyet az [5] dolgozat szerzői egy sorbakapcsolt tározórendszer tervezésére szolgáló sztochasztikus programozási modell szemléltetésére dolgoztak ki. Most a két sorbakapcsolt tározó közül csak az elsőt tekintjük három egymást követő perióduson át (június, július és augusztus). Feltesszük, hogy a véletlen vízigényeket leíró η_1, η_2, η_3 valószínűségi változók egymástól és a véletlen vízhozamoktól is független, gamma eloszlásúak az alábbi paraméterekkel:

| | várható érték (m^3) | szórás (m^3) | ϑ | λ |
|----------|-------------------------|------------------|---------------|---------------|
| η_1 | 215 760 | 327 120 | 0,000 002 016 | 0,435 038 479 |
| η_2 | 433 608 | 243 600 | 0,000 007 307 | 3,168 400 000 |
| η_3 | 484 416 | 214 368 | 0,000 010 541 | 5,106 426 041 |

Hasonlóan feltesszük, hogy a véletlen vízhozamokat leíró ξ_1, ξ_2, ξ_3 valószínűségi változók egymástól és a véletlen vízigényektől is független, gamma eloszlásúak az alábbi paraméterekkel:

| | várható érték (m^3) | szórás (m^3) | ϑ | λ |
|---------|-------------------------|------------------|---------------|---------------|
| ξ_1 | 464 822 | 186 984 | 0,000 013 295 | 6,179 658 245 |
| ξ_2 | 320 576 | 266 040 | 0,000 004 529 | 1,452 005 071 |
| ξ_3 | 2660 40 | 234 040 | 0,000 004 857 | 1,292 152 284 |

A vízkivételre szolgáló készülék forintban mért ára legyen a következő szakaszonként lineáris függvény:

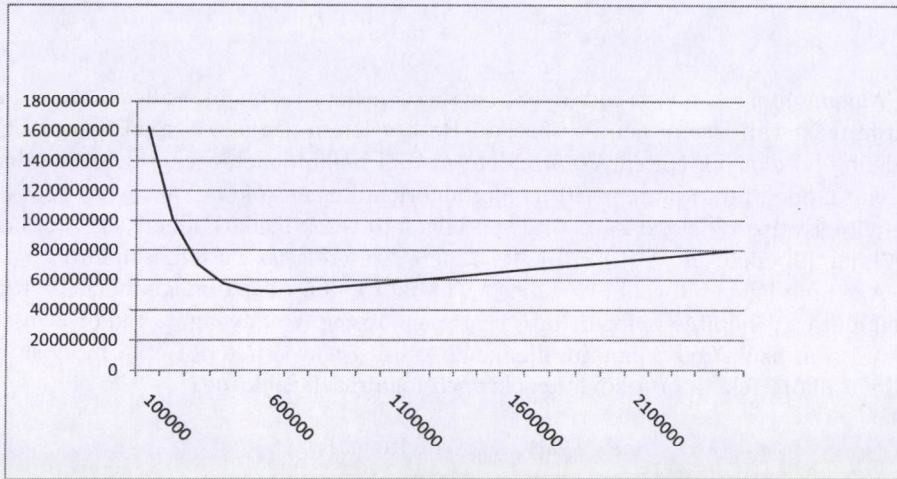
$$p(m) = \begin{cases} 100m, & \text{ha } m \leq 500\,000, \\ 50\,000\,000 + 150(m - 500\,000), & \text{ha } m > 500\,000, \end{cases}$$

és tegyük fel, hogy 25 000 000 m^3 -nél nagyobb kapacitású vízkivételi művet nem építhetünk ki.

Egy m^3 öntözővíz haszna az egyes periódusokban legyen rendre $c_1 = 200$ Ft, $c_2 = 300$ Ft, $c_3 = 250$ Ft. Legyen $N = 10$ és a konstans inflációs ráta $p = 0,05$.

Ekkor a (7) egyváltozós optimalizálási feladat a Matlab programrendszer néhány beépített függvénye (gamma, gammaln, quad, dblquad, fminbnd) segítségével könnyen megoldható.

A fent leírt tesztfeladat optimális megoldása $m = 580\,391\,m^3$ és a hozzá tartozó optimum érték 523 146 000 Ft lett. Az 1. ábra a (7) optimalizálási feladat célfüggvény értékeit mutatja grafikusán az m változó megengedett értékeinek teljes tartományára.



1. ábra. A (7) optimalizálási feladat célfüggvény értékeinek grafikonja.

4. Optimalizálási modell a tározható víz esetére

Egymás utáni periódusokat tekintünk és bevezetjük az alábbi jelöléseket:

| | |
|----------------------|--|
| ξ_0 | a tározóban lévő víz mennyisége az első periódus elején, |
| ξ_k | a vízhozam a k -adik periódusban, |
| $a_k(b_k)$ | a tározóban lévő víz megengedett legkisebb (legnagyobb) mennyisége a k -adik periódusban, |
| z_k | az öntözésre használandó víz mennyisége a k -adik periódusban, |
| N | a periódusok száma, |
| $f(z_1, \dots, z_N)$ | az öntözővíz hasznának jelenértéke, ha az egyes periódusokban z_1, \dots, z_N vízmennyiséget használtunk el, |
| m | a vízkivételt szolgáló műtárgy kapacitása, |
| p | általunk előírt, 1-hez közeli megbízhatósági szint. |

A modellt alkotó feladatot a következőképpen fogalmazzuk meg:

$$\begin{aligned} & \max[f(z_1, \dots, z_N) - p(m)] \quad \text{feltéve, hogy} \\ & P \left\{ a_k \leq \xi_0 + \sum_{j=1}^k \xi_j - \sum_{j=1}^k z_j \leq b_k, \quad k = 1, \dots, N \right\} \geq p \\ & 0 \leq z_k \leq m, \quad k = 1, \dots, N. \end{aligned} \quad (9)$$

Amennyiben m adott, akkor ezt egyszerűen nem tekintjük változónak, a feladaton mást változtatni nem szükséges. Ha a véletlen nagyságú η_k vízigényeket a modellbe be akarjuk építeni, ez minden további nélkül lehetséges a 3. szakaszban tárgyalt modell mintájára. A (9) feladat algoritmikus megoldására a ξ_1, \dots, ξ_N valószínűségi változók eloszlására tett speciális feltevések mellett lehetőség van, lásd az [7], [8], [9] cikkeket. A (9) modellt m méretezésére célszerű felhasználni.

A (9) modell egy további variánsa az, amikor a döntéshozó megadhatja a vízkivételi műtárgy kiépítésére fordítható $p(m)$ pénzösszeg K nagyságát. Ekkor $p(m)$ -et kivonni sem szükséges a maximalizálandó célfüggvény értékéből és a módosított modellt alkotó feladatot a következőképpen fogalmazhatjuk meg:

$$\begin{aligned} & \max f(z_1, \dots, z_N) \quad \text{feltéve, hogy} \\ & P \left\{ a_k \leq \xi_0 + \sum_{j=1}^k \xi_j - \sum_{j=1}^k z_j \leq b_k, \quad k = 1, \dots, N \right\} \geq p \\ & p(m) \leq K 0 \leq z_k \leq m, \quad k = 1, \dots, N. \end{aligned} \quad (10)$$

Érdemes megemlíteni, hogy ha a ξ_1, \dots, ξ_N valószínűségi változók együttes eloszlása folytonos és sűrűségfüggvénye logaritmikusan konkáv, akkor a (9) és (10) feladatok (m, z_1, \dots, z_N) megengedett megoldásainak halmaza konvex (lásd például Prékopa [6]). Ha tehát $f(z_1, \dots, z_N)$ és $p(m)$ konvex függvények, akkor a (9) és (10) feladatok konvexek.

Tekintsünk a (10) feladatra példaként egy tározót négy egymást követő hónapon át, mondjuk április elejétől július végéig, amely havi vízhozam adatai együttes normális eloszlásúak az alábbi várható értékekkel, szórásokkal és korrelációkkal:

| | várható érték ($10^6 m^3$) | szórás ($10^6 m^3$) | korrelációs együtthatók | | | |
|---------|------------------------------|-----------------------|-------------------------|-------|--------|-------|
| ξ_1 | 79,74 | 83,51 | 1,000 | 0,284 | -0,017 | 0,047 |
| ξ_2 | 29,78 | 63,11 | 0,284 | 1,000 | 0,333 | 0,198 |
| ξ_3 | -4,52 | 73,98 | -0,017 | 0,333 | 1,000 | 0,579 |
| ξ_4 | -43,44 | 73,96 | 0,047 | 0,198 | 0,579 | 1,000 |

Képezzük ezekből az aggregált

$$\zeta_1 = \xi_1$$

$$\zeta_2 = \xi_1 + \xi_2$$

$$\zeta_3 = \xi_1 + \xi_2 + \xi_3$$

$$\zeta_4 = \xi_1 + \xi_2 + \xi_3 + \xi_4$$

valószínűségi változókat.

Ezek, mint a $\xi_1, \xi_2, \xi_3, \xi_4$ valószínűségi változók lineáris transzformáltjai, továbbra is együttes normális eloszlásúak az alábbi transzformált várható értékekkel, szórásokkal és korrelációkkal:

| | várható érték ($10^6 m^3$) | szórás ($10^6 m^3$) | korrelációs együtthatók | | | |
|-----------|---------------------------------|--------------------------|----------------------------|-----------|-----------|-----------|
| ζ_1 | 79,740 | 83,510 | 1,000 000 | 0,858 792 | 0,670 483 | 0,542 108 |
| ζ_2 | 109,520 | 118,112 | 0,858 792 | 1,000 000 | 0,872 681 | 0,735 707 |
| ζ_3 | 105,000 | 149,408 | 0,670 483 | 0,872 681 | 1,000 000 | 0,934 830 |
| ζ_4 | 61,560 | 191,201 | 0,542 108 | 0,735 707 | 0,934 830 | 1,000 000 |

Tegyük fel, hogy a (10) optimalizálási feladatban $f(z_1, z_2, z_3, z_4) = 40z_1 + 70z_2 + 80z_3 + 50z_4$, azaz az öntözővíz haszna a felhasznált vízmennyiségek lineáris függvénye. Legyen az m kapacitású vízkivételt szolgáló műtárgy telepítési költsége is lineáris függvény: $p(m) = 50m$. Legyen a tározó vízszintjének legkisebb értékére minden periódusban $a_k = 100, k = 1, 2, 3, 4$; legnagyobb értékére minden periódusban $b_k = 1000, k = 1, 2, 3, 4$ előírva és tegyük fel, hogy az első periódus kezdetén teli tározóval indul a szezon. Ekkor, ha különböző kiépítési költség korlátok mellett megoldjuk az így keletkező optimalizálási feladatot, akkor a döntéshozó elemezni tudja, hogy mekkora kapacitású vízkivételi műtárgyat érdemes kiépíteni. A valószínűségi korlát valószínűségeen belüli kifejezését új változók bevezetésével kicsit egyszerűsítve a K kiépítési költség korlát különböző értékeire végülis az alábbi optimalizálási feladatot oldottuk meg:

$\max(40z_1 + 70z_2 + 80z_3 + 50z_4)$ feltéve, hogy

$$l_1 = 100 + z_1 - 1000$$

$$l_2 = 100 + z_1 + z_2 - 1000$$

$$l_3 = 100 + z_1 + z_2 + z_3 - 1000$$

$$l_4 = 100 + z_1 + z_2 + z_3 + z_4 - 1000$$

$$u_1 = 1000 + z_1 - 1000$$

$$u_2 = 1000 + z_1 + z_2 - 1000$$

$$u_3 = 1000 + z_1 + z_2 + z_3 - 1000$$

$$u_4 = 1000 + z_1 + z_2 + z_3 + z_4 - 1000$$

$$\begin{aligned}
 100P \left(\begin{array}{cccc} l_1 & \leq & \zeta_1 & \leq & u_1 \\ l_2 & \leq & \zeta_2 & \leq & u_2 \\ l_3 & \leq & \zeta_3 & \leq & u_3 \\ l_4 & \leq & \zeta_4 & \leq & u_4 \end{array} \right) \geq 90,00 \\
 50m \leq K \\
 z_1 \leq m \\
 z_2 \leq m \\
 z_3 \leq m \\
 z_4 \leq m
 \end{aligned}$$

Vegyük észre, hogy a fenti feladatban a valószínűségi korlátot 100-zal felszorozva szerepeltetjük, hogy a feladat numerikusan stabilabban legyen megoldható. Ekkor az egyedüli nehézséget a valószínűség és parciális deriváltjai értékeinek a számítása jelenti. Ehhez célszerű a valószínűséget a következőképpen előállítani:

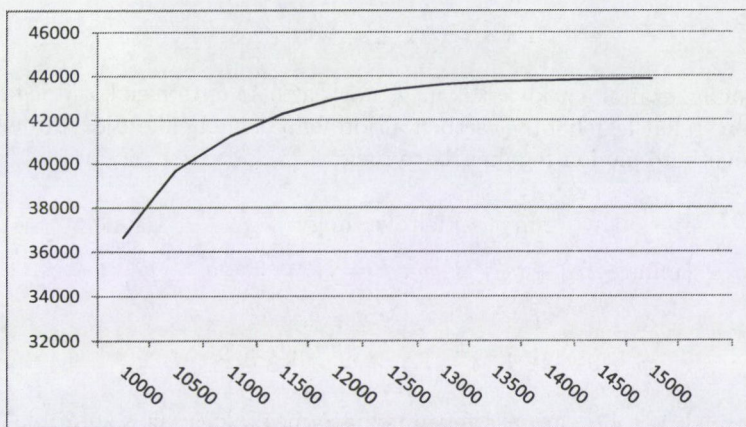
$$\begin{aligned}
 P \left(\begin{array}{cccc} l_1 & \leq & \zeta_1 & \leq & u_1 \\ l_2 & \leq & \zeta_2 & \leq & u_2 \\ l_3 & \leq & \zeta_3 & \leq & u_3 \\ l_4 & \leq & \zeta_4 & \leq & u_4 \end{array} \right) = F(u_1, u_2, u_3, u_4) - F(l_1, u_2, u_3, u_4) \\
 - F(u_1, l_2, u_3, u_4) - F(u_1, u_2, l_3, u_4) \\
 - F(u_1, u_2, u_3, l_4) + F(l_1, l_2, u_3, u_4) \\
 + F(l_1, u_2, l_3, u_4) + F(l_1, u_2, u_3, l_4) \\
 + F(u_1, l_2, l_3, u_4) + F(u_1, l_2, u_3, l_4) \\
 + F(u_1, u_2, l_3, l_4) - F(l_1, l_2, l_3, u_4) \\
 - F(l_1, l_2, u_3, l_4) - F(l_1, u_2, l_3, l_4) \\
 - F(u_1, l_2, l_3, l_4) + F(l_1, l_2, l_3, l_4),
 \end{aligned}$$

ahol $F(x_1, x_2, x_3, x_4)$ a $\zeta_1, \zeta_2, \zeta_3, \zeta_4$ valószínűségi változók megadott paraméterek szerinti együttes normális eloszlásának az eloszlásfüggvényét jelöli. Ez pedig azt jelenti, hogy egyetlen valószínűség értékének a számításához $2^4 = 16$ darab 4-dimenziós normális eloszlásfüggvény értéket kell kiszámítani.

Az így kapott nemlineáris programozási feladatnak elkészítettük az AMPL modelljét, a többdimenziós normális eloszlásfüggvény értékeinek a számításához beillesztettük A. Genz ([1]) numerikus integrálási kódját az AMPL rendszerbe, majd a LOQO solverrel megoldottuk a feladatot a K kiépítési költség korlát különböző értékeire. Az eredményeket a következő táblázat foglalja össze:

| sorszám | K | optimum | m | z_1 | z_2 | z_3 | z_4 |
|---------|-------|-----------|---------|---------|---------|---------|-------|
| 1 | 10000 | 36634,493 | 200,001 | 200,001 | 180,665 | 199,848 | 0,000 |
| 2 | 15000 | 39682,114 | 210,011 | 210,011 | 206,903 | 209,975 | 0,010 |
| 3 | 11000 | 41250,695 | 220,003 | 220,003 | 212,155 | 219,996 | 0,001 |
| 4 | 11500 | 42270,302 | 230,004 | 230,004 | 209,583 | 229,990 | 0,003 |
| 5 | 12000 | 42948,048 | 240,010 | 240,010 | 202,120 | 239,990 | 0,003 |
| 6 | 12500 | 43378,927 | 250,012 | 250,012 | 191,146 | 249,973 | 0,009 |
| 7 | 13000 | 43615,155 | 259,999 | 259,999 | 177,390 | 259,973 | 0,003 |
| 8 | 13500 | 43741,739 | 269,943 | 268,574 | 162,960 | 269,943 | 0,002 |
| 9 | 14000 | 43792,337 | 279,971 | 268,973 | 151,969 | 279,971 | 0,001 |
| 10 | 14500 | 43836,424 | 289,895 | 268,947 | 141,354 | 289,895 | 0,000 |
| 11 | 15000 | 43861,877 | 299,046 | 268,963 | 132,222 | 299,046 | 0,002 |

A K kiépítési költség korlát különböző értékei mellett elérhető öntözési hasznokat a 2. ábra grafikonján is megjelenítettük. Ez a grafikon hasznos lehet a döntéshozó számára annak eldöntésében, hogy meddig érdemes növelni a vízkivételi műtárgy kiépítésére fordítandó pénzösszeget annak függvényében, hogy az mekkora növekedést jelent az öntözővíz hasznosulásában. A döntés meghozatalakor természetesen azt is figyelembe kell venni, hogy a vízkivételi műtárgy kiépítése egyszeri költséget jelent, míg az öntözővíz haszna több éven át realizálható.



2. ábra. Az öntözővíz hasznosulása a vízkivételi műtárgy kiépítésére fordítható pénzösszeg függvényében.

5. Sztochasztikus programozási modell, melyben a vízhiányos napok számát korlátozzuk

Adott időszakot tekintünk, ez lehet pl. az év, mondjuk augusztus hónapja. Előírjuk majd, hogy b -nél több napos kiesés adott, 1-hez közeli p valószínűséggel ne legyen. Modellünket egy n napból álló periódusra írjuk fel. Bevezetjük a következő jelöléseket.

| | |
|-----------------------------|--|
| ξ_1, \dots, ξ_n | napi vízhozamok, |
| $\gamma_1, \dots, \gamma_n$ | napi csapadékok, |
| η_1, \dots, η_n | napi vízigények, |
| m | a vízkivételi műtárgy 1 napi kapacitása, |
| M | m felső korlátja, |
| $p(m)$ | a vízkivételi műtárgy ára. |

A k -adik napon van elegendő öntözővíz akkor és csak akkor, ha fennáll az alábbi összefüggés

$$\min(\xi_k, m) + \gamma_k \geq \eta_k. \quad (11)$$

Legyenek x_1, \dots, x_n csak a 0 és az 1 értékeket felvevő determinisztikus változók. Az alábbi

$$\min(\xi_k, m) + \gamma_k \geq x_k \eta_k \quad (12)$$

összefüggés nem jelent korlátozást, ha $x_k = 0$, de korlátozást jelent, mégpedig nem mást, mint (11)-et, ha $x_k = 1$. Előírva az

$$x_1 + \dots + x_n \geq n - b$$

feltételt (12) mellé, ezáltal azt kívántuk meg, hogy az (11) feltételek közül legalább $n - b$ teljesüljön, tehát legfeljebb b esetben álljon fenn (11) ellenkezője. Modellünk ezek után a következő módon fogalmazható meg:

$$\begin{aligned} &\min p(m) \text{ feltéve, hogy} \\ &P\{\min(\xi_k, m) + \gamma_k \geq x_k \eta_k, \quad k = 1, \dots, n - b\} \geq p \\ &x_1 + \dots + x_n \geq n - b, \\ &x_k = 0 \text{ vagy } 1, \quad k = 1, \dots, n, \quad 0 \leq m \leq M. \end{aligned} \quad (13)$$

Ennek a modellnek is több variánsa fogalmazható meg, akár csak a korábbiaknak. Többek között beépíthető a célfüggvénybe egy b -től függő költségtenyező. Ha a $\xi_k, \eta_k, \gamma_k, k = 1, \dots, N$ valószínűségi változók együttes eloszlása folytonos és az együttes sűrűségfüggvény logaritmikusan konkáv, akkor a (13) feladat feltételei, kivéve az $x_k = 0, 1$ feltételeket, konvex halmazt határoznak meg.

Hivatkozások

- [1] A. GENZ: *Numerical Computation of the Multivariate Normal Probabilities*. Journal of Computational and Graphical Statistics 1 (1992) 141–150.
- [2] B. V. GNYEGYENKO, A. N. KOLMOGOROV: *Független valószínűségi változók összegeinek határeloszlásai*. Akadémiai Kiadó, Budapest (1951).
- [3] A. PRÉKOPA: *On secondary processes generated by random point distributions of Poisson type*. Annales Univ. Sci. R. Eötvös, Sectio Math. 2 (1959) 139–146.
- [4] A. PRÉKOPA: *Contributions to the theory of stochastic programming*. Mathematical Programming 4 (1973) 202–221.
- [5] A. PRÉKOPA, T. RAPCSÁK, I. ZSUFFA: *Egy új módszer sorbakapcsolt tározórendszer tervezésére sztochasztikus programozás felhasználásával*. Alkalmazott Matematikai Lapok, 2 (1976) 189–2001.
- [6] A. PRÉKOPA: *Stochastic Programming*. Kluwer Scientific Publishers Dordrecht, Boston (1995).
- [7] A. PRÉKOPA, T. SZÁNTAI: *On Optimal Regulation of a Storage Level with Application to the Water Level Regulation of a Lake*. European Journal of Operational Research 3 (1979) 175–189.
- [8] T. SZÁNTAI: *Evaluation of a Special Multivariate Gamma Distribution*. Mathematical Programming Study 27 (1986) 1–16.
- [9] T. SZÁNTAI: *A Computer Code for Solution of Probabilistic Constrained Stochastic Programming Problems*. In: Numerical Techniques for Stochastic Optimization (Yu. Ermoliev, R. J.-B. Wets, eds.), Springer, New York, (1988) 229–235.
- [10] L. TAKÁCS: *Poisson folyamat által származtatott másodlagos folyamatokról és azok fizikai alkalmazásairól*. MTA Matematikai és Fizikai Tudományok Osztályának Közleményei 4 (1954) 473–504.

(Beérkezett: 2010. április 24.)

PRÉKOPA ANDRÁS
Rutcor, Rutgers University
prekopa@rutcor.rutgers.edu
ELTE TTK
Operációkutatási Tanszék
prekopa@cs.elte.hu

SZÁNTAI TAMÁS
BME TTK
Matematikai Intézet, Differenciálegyenletek Tanszék
szantai@math.bme.hu

DETERMINATION OF THE IRRIGATION WATER CONTENT OF
NATURAL STREAMFLOWS

A. PRÉKOPA, T. SZÁNTAI AND I. ZSUFFA

In the first part of the paper a theorem is proved which states that if rainfalls occur according to a Poisson process and the inflows have gamma density profiles with independent, exponentially distributed amplitudes, then the stationary distribution of the streamflow intensity is gamma. In the further parts of the paper three optimization models are presented that lead to three different solutions of the problem. In the first one the optimal capacity of a pump station is to be determined when the water cannot be stored. In the second one the optimal capacity of a reservoir, where the water can be stored, and an optimal water usage policy are to be found, given that the water demands should be met by prescribed large probability. In the third one an upper bound is imposed on the number of days when demands may not be met and the cost of a pump station is to be minimized.

A folyók öntözővíz készlete meghatározásának problémáját Zsuffa István vetette fel több évvel ezelőtt. A feladat részletes kidolgozására nemrég került sor az első két szerző részéről, akik ezt a közös dolgozatot barátjuk és munkatársuk, Zsuffa István emlékének ajánlják.

The problem of finding the irrigation water contents of rivers was formulated by István Zsuffa many years ago. The detailed elaboration of the problem is more recent and is due to the first two authors who offer this paper to the memory of their friend and co-worker, István Zsuffa.

TURÁNYI TAMÁS: REAKCIÓMECHANIZMUSOK VIZSGÁLATA

AKADÉMIAI KIADÓ, BUDAPEST, 2010.

TÓTH JÁNOS

Mi indokolja ennek az ismertetésnek a megjelentetését az *Alkalmazott Matematikai Lapokban*?

Turányi Tamás vegyész és alkalmazott matematikus hosszú évek óta foglalkozik nagy összetett kémiai reakciók modellezésével, egyszerűsítésével és numerikus vizsgálatával. Az itt szereplő modellek leggyakrabban nemlineáris (sokszor polinomiális) közönséges, ritkábban parciális differenciálegyenletről álló rendszerek. A *nagy* jelző arra utal, hogy az egyenletek száma, vagyis a figyelemmel kísért anyagfajták száma legalább néhány tucat, de esetenként többszáz, sőt több ezer is lehet. Hogyan foghatunk hozzá egy ilyen, matematikusi szemmel elrettentő rendszer vizsgálatához?

A szerző a legfontosabb alapfogalmak és alapproblémák ismertetése (2. *Néhány reakciókinetikai alapismeret*) után a következő megközelítéseket ajánlja.

3. *Reakcióutak* A reakcióutak igen szemléletes heurisztikus módszere Horiutól [10] és Temkintől (Tyomkin) [17] származik. A könyv néhány ábrája ennek a módszernek a felhasználásával megmutatja egyes elemek átvitelét a különféle anyagfajták között metán-levegő elegyek robbanása közben. A módszer részletes és matematikailag kielégítő tárgyalása nem ismeretes, feltehetőleg az operációkutatás, speciálisan a hálózati folyamatok elmélete nyújtana ehhez alkalmas eszközöket.

4. *Érzékenység- és bizonytalanságanalízis* A közönséges differenciálegyenletek elméletében – az elsőrendű lineáris parciális differenciálegyenletekkel való kapcsolat megvilágítása végett – elő szokott kerülni a *variációs egyenlet*. Ugyanez az egyenlet a műszaki és a reakciókinetikai irodalomban *érzékenységi egyenlet*ként szerepel, ugyanis a megoldások paraméterek szerinti deriváltja úgy is interpretálható, mint a megoldások érzékenysége a paraméterek megváltoztatására, ezért is nevezik őket ezen a tájon *lokális érzékenysége*eknek vagy *lokális érzékenységi együttható*knak. Mivel meghatározásuk számításgényes feladat, ezért folyamatosan újabb és újabb heurisztikus eljárásokat dolgoznak ki erre a célra. Megjegyzendő, hogy a reakciókinetikán kívüli érzékenységvizsgálattal foglalkozó irodalom jelentős része sokkal egyszerűbb objektumokkal, időtől független megoldásfüggvények paramétertől való függésével foglalkozik.

A bizonytalanságanalízis lényege, hogy a modellek paramétereit valószínűségi változóknak tekintve képesek vagyunk-e valamit mondani a modellek felhasználásával számolt numerikus eredmények eloszlásáról. Ennek a fejezetnek különösen

szép része a reakciósebességi együttható hőmérsékletfüggését leíró Arrhenius-összefüggés paraméterei bizonytalanságának elemzése.

5. *Időskála-analízis* Itt nem a Stefan Hilger által kezdeményezett [8] és újabban egyre nagyobb érdeklődést kiváltó [5, 19] időskálakon értelmezett dinamikáról van szó, hanem a szinguláris perturbáció klasszikus, Tyihonov-féle elméletének [14, 18, 20] alkalmazásáról. Fizikai folyamatoknál és kémiai reakcióknál ugyanis meglehetősen tipikus, hogy a folyamatok több időskálán zajlanak, egyesek sokkal gyorsabbak a többiekénél. Az ilyen fizikai és kémiai folyamatokat leíró differenciálegyenletek *merevek* és numerikus megoldásuk nehéz. A merev differenciálegyenletek kezelhetetlensége speciális numerikus módszerek létrehozását kényszerítette ki, amilyen például Gear módszere [6]. A több, egymástól nagyon különböző időskála léte nemcsak hátrányos: az ilyen folyamatok vizsgálata leegyszerűsíthető azáltal, hogy az egyes időskálakon végbemenő folyamatokat külön kezeljük.

Az egyszerűsítés matematikai alapjáról szól egy fontos speciális esetben (a Michaelis–Menten-reakció esetében) Heineken, Tsuchiya és Aris több mint 40 évvel ezelőtti alapvető, még mindig nem elég széles körben ismert munkája [7]. Mivel a téma általános és matematikailag korrekt tárgyalása várat magára, ezért még napjainkban is lehet közölni meglehetősen naiv cikkeket erről a területről. A terület kvázistacionárius közelítés néven a könyv 6.5. szakaszában újra előkerül.

6. *Reakciómechanizmusok redukciója* A nagy rendszerek egyszerűsítésére irányuló törekvések egyik célja a változók számának csökkentése. A szerző itt nemcsak a szűk értelemben vett változóösszevonás (*lumping*) technikáját ismerteti, ami itt anyagfajták összevonása néven szerepel, hanem számos további hasznosnak bizonyult eljárást is, így a felesleges anyagfajták és reakciólépések elhagyását és a reakciólépések összevonását is.

Szellemes és rendkívül hatékony a *repromodellezés* módszere, ami nem más, mint a kinetikai differenciálegyenlet megoldó operátorának interpolálása polinommal, majd ennek felhasználása a reakciódifúzió-egyenlet megoldására. Az eljárást parciális differenciálegyenletek megoldására csak az operátorszeletelés módszerével együtt lehet alkalmazni. Ez utóbbit Magyarországon elsősorban Faragó István és munkatársai vizsgálják [2, 11] és alkalmazzák. A repromodellezés módszerével kapott számítások eredményei teljesen meggyőzőek, a közelítések pontosságára és a szükséges gépidőre vonatkozó általános matematikai tárgyalásról nincs tudomásom.

7. *Az érzékenységi függvények hasonlósága* Ez a terület a legjellemzőbb abból a szempontból, hogy matematikus nem merne hozzáfogni, Turányi Tamás és munkatársai viszont bátran alkalmazzák, mint heurisztikus módszert, és például rendkívül érdekes biológiai tanulságokhoz jutnak el. Bonyolult molekuláris biológiai modellekben egyes paraméterek (rendszerint reakciósebességi együtthatók) szerinti érzékenységek hasonlósága azt jelentheti, hogy valamelyik reakciólépés sebességi együtthatójának jelentős megváltozását más reakciólépések sebességi együtthatójának megfelelő megváltoztatásával kompenzálni lehet. Ez a biológiai rendszerek robosztusságának egyik magyarázata lehet.

8. *Programok összetett reakciómechanizmusok vizsgálatára* A könyvben leírt módszerek többségét nem kell beprogramozni, hanem az Internetről számos

program letölthető az eljárások használatához. Áttekintést kapunk az általános programokról, amelyek kinetikai differenciálegyenletek megoldására vagy a folytonos idejű, diszkrét állapotterű sztochasztikus modell (lásd például Rényi cikkét 1953-ból: [15]) szimulálására [16], illetve elemzésére szolgálnak. Az általános matematikai módszerek mellett a szerző – érdeklődésének megfelelően – ismerteti az égési, légkörkémi és biokémiai kinetikai modelleket kezelő speciális szimulációs és analízis programokat is.

9. *Összefoglalás* Ez a könyv záró fejezetének címe, de itt a recenzió összegezése következik.

A könyv nem kis részben a szerző, kollégái és tanítványai nemzetközi folyóiratokban megjelent munkáin alapul. Ugyanakkor a könyv több témakör széleskörű áttekintését is tartalmazza és különösen hasznos az elképesztően gazdag (464 tételt tartalmazó) irodalomjegyzék. Az Internet pozitív hatása több helyen is érezhető. A könyv számos számítógépes program honlapjának címét is megadja. A könyvvel kapcsolatos újabb híreket és kiegészítéseket a

<http://garfield.chem.elte.hu/Turanyi/reakciomechanizmusok.html>

weboldalon lehet elolvasni, és a könyv néhány kedvcsináló fejezetét is le lehet onnan tölteni.

Igen hasznos a területtel ismerkedő számára, hogy a *Tárgymutató* a kifejezések angol eredetijét is megadja. A könyv tartalomjegyzéke angol nyelven is szerepel, remélhetőleg felkeltve valamely nemzetközi kiadó érdeklődését is.

Ami a formát illeti: a tipográfia megfelelő, a megértéshez jelentősen hozzájárul az ábrák szépek, a képletek persze olyanok, amilyeneket a Word megenged...

Egy kritikai megjegyzés: ha egy ilyen könyv irodalomjegyzékéből Craciun [1], Feinberg [3, 4], Horn és Mincheva [12] neve hiányzik (csak Volpert képviseli a formális reakciókinetika matematikai elméletét), akkor helyesebb lett volna címként ezt adni: *Reakciómechanizmusok számítógépes vizsgálata*, mert a legnagyobb hangsúly a számítógépes módszerekre esik.

A mű a szerző szándéka szerint segítséget kíván nyújtani egyetemi hallgatóknak, illetve a területtel ismerkedni kívánó kutatóknak. Ezen feladatainak a könyv kiválóan megfelel. A jelen ismertetést olvasó alkalmazott matematikusok számára azonban a fentiekből az is kiderülhetett, hogy a reakciókinetikai alkalmazások területén nem kevesebb és nem könnyebb feladatok várnak a – statisztikával, analízissel, numerikus matematikával foglalkozó – matematikusokra (is), mint a rozsomákok vándorlásának területén [13], és talán ezeknek a feladatoknak a megoldása még némi (egyéni anyagi és közérdekű környezetvédelmi) haszonnal is járhat.

Hivatkozások

- [1] CRACIUN, M., FEINBERG, M.: *Multiple equilibria in complex chemical reaction networks: II. The species-reaction graph*. SIAM J. Appl. Math. **66**, (4) (2006) 1321–1338.
- [2] FARAGÓ, I., HAVASI, Á.: *Consistency analysis of operator splitting methods for $C0$ -semigroups*. Semigroup Forum **74**, (2007) 125–139.

- [3] FEINBERG, M.: *Chemical reaction network structure and the stability of complex isothermal reactors: I. The deficiency zero and deficiency one theorems*. Chem. Eng. Sci. **42**, (10) (1988) 2229–2268.
- [4] FEINBERG, M.: *Chemical reaction network structure and the stability of complex isothermal reactors: II. Multiple steady states for networks of deficiency one*. Chem. Eng. Sci. **43**, (1) (1988) 1–25.
- [5] GARAY, B. M., VÁRDAI, J.: *Interpolation of dynamic equations on time scales*. J. Difference Eq. Appl. **13**, (8–9) (2007) 847–854.
- [6] GEAR, C. W.: *The automatic integration of ordinary differential equations*. Communications of the ACM **14**, (3) (1971) 176–179.
- [7] HEINEKEN, F. G., TSUCHIYA, H. M., ARIS, R.: *On the mathematical status of the pseudo-steady state hypothesis of biochemical kinetics*. Math. Biosci. **1**, (1) (1967) 95–113.
- [8] HILGER, S.: *„Ein Maßkettenkalkül mit Anwendung auf Zentrums Mannigfaltigkeiten“*, Thesis, Universität Würzburg, (1998).
- [9] Honlap a szinguláris perturbációról és alkalmazásairól:
<http://www.ima.umn.edu/~milik/singdir.html#apl:chem>
- [10] HORIUTI, J.: *Theory of reaction rates as based on the stoichiometric number concept*. Annals of the New York Academy of Sciences **213**, (1973) 5–30.
- [11] LADICS, T.: *The analysis of the splitting error for advection-reaction problems in air pollution models*. Időjárás – (Quarterly Journal of the Hungarian Meteorological Service) **109**, (3) (2005) 173–188.
- [12] MINCHEVA, M., SIEGEL, D.: *Tömeghatás típusú reakciódiffúzió-rendszerek stabilitása*. Alk. Mat. Lapok **26**, (2009) 97–127. (Fordította: Egri Edit).
- [13] Rozsomákokról:
http://bme.ysolt.net/3_felev/Matek_A4/matb4gyakorlat_2005_mo.pdf
- [14] O'MALLEY, R. E., JR.: *Singular perturbation methods for ordinary differential equations* (Springer-Verlag, New York, 1991).
- [15] RÉNYI, A.: *Kémiai reakciók tárgyalása a sztochasztikus folyamatok elmélete segítségével*. Magy. Tud. Akad. Mat. Kut. Int. Közl. **2**, (1953) 83–101.
- [16] SIPOS, T., TÓTH, J., ÉRDI, P.: *Stochastic simulation of chemical reaction by digital computer, I. The model*. React. Kinet. Catal. Lett. **1**, (1) (1974) 113–117.
- [17] TEMKIN, O., ZEIGARNIK, A., BONCHEV, D.: *Chemical Reaction Networks: A Graph Theoretical Approach*, (CRC Press, 1996).
- [18] VASIL'eva, A. B.: *Andrey Nikolaevich Tikhonov and his school for singular perturbation problem*. Matem. Mod. **13**, (12) (2001) 6–9.
- [19] YANTIR, A., ÜNAL UFUKTEPE, Ü.: *Mathematica applications on time scales*. Lecture Notes in Computer Science **3482**, (2005) 529–537.
- [20] ZACHÁR, A.: *Comparison of transformations from nonkinetic to kinetic models*. Acta Chimica Hungarica – Models in Chemistry **135**, (3) (1998) 425–434.

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

Alkalmazott Matematikai Lapok (2010)

MTA Könyvtára
Permalink 20100628... n.

2011 OKT 25

Az Alkalmazott Matematikai Lapok megjelenését támogatja
a Magyar Tudományos Akadémia Könyv- és Folyóiratkiadó Bizottsága.

A kiadásért felelős a BJMT főtítkára
Szedte és tördelte Éliás Mariann

Nyomta a Nagy és Társa Kft., Budapest

Budapest, 2010
Megjelent 18 (A/5) ív terjedelemben
250 példányban
HU ISSN 0133-3399

TÓTH JÁNOS

BME, TTK

Analízis Tanszék

Budapest, Egry J. u. 1.

jtoth@math.bme.hu

A jelen ismertető már útban volt a nyomda felé, amikor megjelent a hír, hogy a Kémiai Tudományok Osztályának ajánlása alapján az Akadémiai Kiadó Nívódíját vehette át Maróth Miklóstól, az MTA alelnökétől Turányi Tamás, az MTA doktora, a Reakciómechanizmusok vizsgálata című könyvéért.

ÚTMUTATÁS A SZERZŐKNEK

Az Alkalmazott Matematikai Lapok csak magyar nyelvű dolgozatokat közöl. A közlésre szánt dolgozatokat e-mailen az `aml@math.elte.hu` címre kérjük elküldeni az ábrákat tartalmazó fájlokkal együtt. Előnyben részesülnek a \LaTeX -ben elkészített dolgozatok.

A kéziratok szerkezeti felépítésének a következő követelményeket kell kielégíteni:

Fejléc: A fejlécnek tartalmaznia kell a dolgozat címét és a szerző teljes nevét.

Kivonat: A fejléc után egy, képletet nem tartalmazó, legfeljebb 200 szóból álló kivonatot kell minden esetben megadni.

Fejezetek: A dolgozatot címmel ellátott szakaszokra kell bontani, és az egyes szakaszokat arab sorszámozással kell ellátni. Az esetleges bevezetésnek mindig az első szakaszt kell megnevezni.

A dolgozatban előforduló képleteket a dolgozat szakaszokra bontásától független, folytatólagos arab sorszámozással kell azonosítani. Természetesen nem szükséges minden képletet számozással ellátni, csak azokat, amelyekre a szerző a dolgozatban hivatkozni kíván.

Mind az ábrákat, mind a lábjegyzeteket szintén folytatólagos arab sorszámozással kell ellátni. Az ábrák elhelyezését a dolgozat megfelelő helyén ábraazonosító sorszámokkal kell megadni. A lábjegyzetekre a dolgozaton belül az azonosító sorszám felső indexkénti használatával lehet hivatkozni.

Az esetleges definíciókat és tételeket (segéd tételeket és lemmákat) szakaszonként újrakezdődő, ponttal elválasztott, kettős számozással kell ellátni. Kérjük a szerzőket, hogy ezeket, valamint a tételek bizonyítását a szövegben kellő módon emeljék ki.

Irodalomjegyzék: A dolgozatok szövegében az irodalmi hivatkozás számait szögletes zárójelben kell megadni, mint például [2] vagy [1, 7–13].

Az irodalmi hivatkozások formája a következő: Minden hivatkozást fel kell sorolni a dolgozat végén található irodalomjegyzékben, a szerzők, illetve a társszerzők esetén az első szerző neve szerint alfabetikus sorrendben úgy, hogy a cirill betűs szerzők nevét a Mathematical Reviews átirási szabályai szerint latin betűsre kell átirni. A folyóiratban megjelent cikkekre [1], a könyvekre [2] a következő minta szerint kell hivatkozni:

[1] FARKAS, J.: *Über die Theorie der einfachen Ungleichungen*, Journal für die reine und angewandte Mathematik 124, (1902) 1–27.

[2] ZOUTENDIJK, G.: *Methods of Feasible Directions*, Elsevier Publishing Company, Amsterdam and New York (1960), 120 o.

Szerző adatai: Az irodalomjegyzék után, a kézirat befejezéseképpen fel kell tüntetni a szerző teljes nevét és a munkahelye (esetleg lakása) pontos címét, illetve e-mail címét.

Idegen nyelvű kivonat: Minden dolgozathoz csatolni kell egy angol nyelvű összefoglalót.

A szerzők a dolgozatukról 20 darab ingyenes különlenyomatot kapnak. A dolgozatok után szerzői díjat az Alkalmazott Matematikai Lapok nem fizet.

317.471

27/2010

TARTALOMJEGYZÉK

| | |
|--|-----|
| <i>Szathmáry László és Bognár Katalin, ZART: egy multifunkcionális mintakereső algoritmus</i> | 107 |
| <i>Benkő Attila, Dósa György, Egy új feladat: ládafedés szállítással és ennek megoldása algoritmusok evolúciójával</i> | 123 |
| <i>Fazekas István, Karácsony Zsolt, Libor Józsefné, A leghosszabb szériák vizsgálata</i> | 135 |
| <i>Takács Szabolcs, Egy nem hagyományos statisztikai eljárás bemutatása az OECD PISA adatbázison - esettanulmány</i> | 157 |
| <i>Prékopa András, Szántai Tamás, Zsuffa István, Természetes vízfolyások öntözővíz készletének meghatározását célzó matematikai modellek</i> | 175 |
| <i>Tóth János, Könyvismertetés: Turányi Tamás: Reakciómechanizmusok vizsgálata</i> | 189 |

INDEX

| | |
|--|-----|
| <i>László Szathmáry and Katalin Bognár, ZART: a multifunctional itemset mining algorithm</i> | 107 |
| <i>Attila Benkő, György Dósa, Bin covering with delivery and solving it with evolution of algorithms</i> | 123 |
| <i>István Fazekas, Zsolt Karácsony and Józsefné Libor, On longest runs</i> | 135 |
| <i>Szabolcs Takács, A non-traditional statistical method in the OECD PISA database - case study</i> | 157 |
| <i>András Prékopa, Tamás Szántai and István Zsuffa, Determination of the irrigation water content of natural streamflows</i> | 175 |
| <i>János Tóth, Book review</i> | 189 |